

OM nucleic - nucleic search, using sw model

Run on: August 19, 2002, 15:01:07 ; Search time 2294.29 Seconds
(without alignments)
16536.624 Million cell

updates/sec

Title: US-09-902-759-38

Perfect score: 1813

Sequence: 1 ggagccgcctgggtgcag.....cataatgttgratgaaaaa 1813

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1813	100.0	1813	6	AX076924 Sequence
2	1809.8	99.8	1838	9	AF361746 Homo sapi
3	1809	99.8	1821	6	AX136161 Sequence
4	1804.8	99.5	1816	6	AX191598 Sequence
5	1760.4	97.1	1831	6	AX073678 Sequence
6	1691.8	93.3	1734	9	BC016868 Homo sapi
7	1659.6	91.5	1855	9	AB060855 Macaca fa
8	1171.4	64.6	1173	6	AX191588 Sequence
9	861.6	47.5	1840	10	AF361882 Mus muscu
10	828.2	45.7	187960	9	AP000866 Homo sapi
11	793.2	43.8	101458	2	AP000680 Homo sapi
12	717.4	39.6	736	9	AF277292 Homo sapi
13	531.2	29.3	637	6	AX136493 Sequence
14	453	25.0	541	6	AX136640 Sequence
15	441	24.3	441	6	AX0732845 Sequence
16	339	18.7	221961	10	AC073435 Mus muscu
17	323.8	17.9	340	6	AX331694 Sequence
18	323.8	17.9	340	6	AX333904 Sequence
19	183.2	10.1	81318	2	AC016125 Homo sapi
20	183.2	10.1	101458	2	AP000680 Homo sapi
21	162	8.9	187960	9	AP000866 Homo sapi
22	132.8	7.3	637	9	HS3326182
23	132.8	7.3	674	9	HS3329044
24	120	6.6	698	9	HS3331517
25	79.8	4.4	46240	2	AC105958 Mus muscu
26	75	4.1	81318	2	AC016125 Homo sapi
27	71	3.9	674	9	HS3331520
28	70.8	3.9	125020	9	AF429315 Homo sapi
29	70	3.9	611	9	HS3323034
30	59.6	3.3	165556	2	AC087190 Homo sapi
31	58.2	3.2	1147	9	BC007313 Homo sapi
32	58.2	3.2	1161	6	AX056679 Sequence
33	58.2	3.2	1221	9	AK000460 Homo sapi
34	58	3.2	125020	9	AF429315 Homo sapi
35	54.2	3.0	95209	2	AP004323 Oryza sat
36	54.2	3.0	153292	2	AP003635 Oryza sat
37	51.4	2.8	7218	6	I66494 Sequence 14
38	51	2.8	51	6	AX161199 Sequence
39	51	2.8	51	6	AX161201 Sequence
40	50.4	2.8	57245	2	AC068263 Homo sapi
41	50	2.8	50	6	AX076928 Sequence
42	49.6	2.7	144973	2	AC096689 Oryza sat

ALIGNMENTS

RESULT 1

AX076924

LOCUS

AX076924

1813 bp

DNA

linear

PAT

22-FEB-2001

DEFINITION

Sequence 36 from Patent WO0105836.

ACCESSION

AX076924

VERSION

AX076924.1 GI:13121579

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1813)

Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and

Wood,W.I.

Polypeptidic compositions and methods for the treatment of tumors

JOURNAL

Patent: WO 0105836-A 36 25-JAN-2001;

Genentech, Inc. (US)

FEATURES

source

1..1813

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 368 a 559 c 484 g 402 t

ORIGIN

Query Match 100.0%; Score 1813; DB 6; Length 1813;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1813; Conservative 0; Mismatches 0; Indels 0; Gaps

0;

QY 1

ggagcgccctggggtgcaaggctcggctccgcgcacgcgtcggcgcgtagcgcct 60

DB 1

GGAGCGCCCTGGGTGTCAAGGCTCGGCTCCCGGCACAGCTCGGCGCTCGCGCAGCT 60

QY 61

cggcaactgcaggtcggtggtccgcgggtggtggtggtggtggtggtggtggtggt 120

DB 61

CGGCACTTCAGAGTGTGGTGTCCCGGGCTCGCGCCCTGTACTCGTCCGCGCAGGGA 120

QY 121

gggcatgatttcttcctcccggggcccctgggtgaccaacttgctggggtttttgttct 180

DB 121

GGGCCATGATTTCTTCCCGGGGCCCCCTGTGTGACCAACTTGTCTGCGGTTTTGTCTGG 180

QY 181

ggctgagtgccctcgcccccttcctcggggcagctgcaactgcaactgcccgaacc 240

DB 181

GGCTGATGTGCTTCGGGCCCCCTTCGGGGGCCAGGTGCACTGTGCACTTGGCCGCAACC 240

QY 241

gggtcaggcggtgagggagggaagtggtgttcctcagcggtggtacacctgtgacgggg 300

DB 241

GGTTGAGCGCGGTGAGGAGGGAAGTGGTGTCTCCAGCTGTGTACACTTGCACGGG 300

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

181 ggctgaagtcgcttcgagcccccctcgagggccagctgcaactgcaacttgccttcgcccgaacc 240
194 Gbctgagtgcccttcgagccccccttcgagggccagctgcaactgcaacttgccttcgcccgaacc 253
241 ggttgacgagcggtgagggagggaggtggtgcttcagcggtggtacaccttgacacgggg 300
254 ggttgacgagcggtgagggagggaggtggtgcttcagcggtggtacaccttgacacgggg 313
301 aggtgtcttcatccacgacgcatggaggtgccccttggatggtggttcttcaaacagaaag 360
314 aggtgtcttcatccacgacgcatggaggtgccccttggatggtggttcttcaaacagaaag 373
361 aaaggaggatcaggtgtgtcctcatcatcaatgggggtcacacaagcaaacctggagat 420
374 aaaggaggatcaggtgtgtcctcatcatcaatgggggtcacacaagcaaacctggagat 433
421 cettgggtctactccatgctcccgaaacctgcccctgggtggaggggtctccagaga 480
434 ccttggtctactccatgctcccgaaacctgcccctgggtggaggggtctccagaga 493
481 aagactctggccctacagctgctcggtgaatgacagacaaacaggaacaaactaggg 540
494 aagactctggccctacagctgctcggtgaatgacagacaaacaggaacaaactaggg 553
541 gccacagcatcaaaacctagaactcaatgtactggttctccagctcctccatcctgcc 600
554 gccacagcatcaaaacctagaactcaatgtactggttctccagctcctccatcctgcc 613
601 gtctccagggtgagcccatgaggggggcaaacgtgacccctgagctgcccagctccaaagga 660
614 gtctccagggtgagcccatgaggggggcaaacgtgacccctgagctgcccagctccaaagga 673
661 gtaagcccgctgtccaatcaacagtgaggtcggcagcttccatcctccagacttctttg 720
674 gtaagcccgctgtccaatcaacagtgaggtcggcagcttccatcctccagacttctttg 733
721 caccagcatagatgcatcggtgggtcttcaagctcaccacacttggcttccatgg 780
734 caccagcatagatgcatcggtgggtcttcaagctcaccacacttggcttccatgg 793
781 ctggagctatgctcgaaggcccaaatgaggtgggcaactgcccgaatgtaagtgaagc 840
794 ctggagctatgctcgaaggcccaaatgaggtgggcaactgcccgaatgtaagtgaagc 853
841 tggagagtgagcacagggcctggagctgaggtggtgctggaggtggtggtggtacacctgg 900
854 tggagagtgagcacagggcctggagctgaggtggtgctggaggtggtggtggtacacctgg 913
901 tggagagtgagcacagggcctggagctgaggtggtgctggaggtggtggtggtacacctgg 960
914 tggagagtgagcacagggcctggagctgaggtggtgctggaggtggtggtggtacacctgg 973
961 agggacagccaatgatatcaaggaggatgcatgctcccgagacctgacctggccca 1020
974 agggacagccaatgatatcaaggaggatgcatgctcccgagacctgacctggccca 1033

RESULT 3

1034 AGAGCTCAGACACAACTCTCCAAAGATGGAGCCCTTCTCTGTCTACCTCCGACAGGCC 1093
1081 tccggccaccccatggccctcccgaggtcgtgcatgtgaccccaacgcccagctctcca 1140
1094 TCGGCCACCCCATGGCCCTCCGAGGCTGGTGCAATTGACCCCAAGCCGCACTCTCCA 1153
1141 gccaggccctgcccctcaccagaactgcccacagacagatggggggccacccctcaaccaat 1200
1154 GCCAGGCCCTGGCCCTTACCAGACTGCCCAACAGATGGGGGCCACACCTCAACCAATAT 1213
1201 ccccatccctcgtgggggttctcctcagcttgagcgcatgggtcgtgctgctgctgga 1260
1214 CCCCATCCCTGGTGGGGTTCTTCTCTGTGCTTGAGCGCATGGGTGCTGCTCTGTGA 1273
1261 tgggtcctgccagagtcgaagctggctctggtatgatgacdcacacactcatggcta 1320
1274 TGNATCCCTGCCAGAGTCAGCTGGCTCTCTGGTATGATGATGACCCACCACTCATTTGCTA 1333
1321 aaggatttgggggtctcctcctcctataaagggtcacctctagcacagaggtcgtgactg 1380
1334 AAGGATTGGGGTCTCTCTCTTAAGGCTCACCTCTAGCACAGAGGCTTGAGTCATG 1393
1381 ggaagagtcacactcctgacaccttagtactctgccccacactctcttactgtgggaaa 1440
1394 GGAAGAGTCACACTCTTGACCTTAACTTCTGTGCCCCCACCCTCTTTTACTGTGGGAAA 1453
1441 accatctcaatgaagacctcaaggtccagagacagagaagagagagagagagagagagag 1500
1454 ACCATCTCAGTAAGACTTAAGTGTCCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
1501 attgggaggagcctccaccacccctgactcctccttatgaagccagctgctgaaatag 1560
1514 ATTGGGAGGAGCTTCCACCCACCTGACTCTCTCTTATGAGCCAGCTGCTGDAATAG 1573
1561 ctactcaccagaagtgagggggcagagacttcaggtcaetgagctcccgagggcccccctga 1620
1574 CTACTCACCAGAGTGAGGGGAGAGACTTCCAGTCACTGAGTCTCCAGGCCCCCTTGA 1633
1621 tctgtaccaccacctatctaaccaccccttgggtcccatccagctccctcgtatgat 1680
1634 TCTGTACCCCAACCCCTATCTAAACACCCCTTGGCTCCCACTCCAGCTCCTCTGATTTGAT 1693
1681 ataacctcaggtcaggtcgttgggttaggtttactggggcagagagatgggaatccttat 1740
1694 ATAACTCTCAGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1753
1741 taaaactaacatgaatatggtgtgtttcttcttcttcttcttcttcttcttcttcttcttct 1800
1754 TAAACTAATCAATGAATGATGTTGTTTTCATTTTGCAAAATTTAATAAAGATACATATG 1813
1801 ttgttatgaaaaa 1813
1814 TTGTATGAAAAA 1826

[illegible]

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: http://image.llnl.gov Series: IRAK Plate: 20 Row: m Column: 10.

FEATURES
Location/Qualifiers
source 1..1734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:17599 IMAGE:3850665"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_65"
/lab_host="DH10B"
/note="vector: pcwv-SPORT6"
CDS
8..1180
/codon_start=1
/product="unknown (protein for MGC:17599)"
/protein_id="AAH16868.1"
/db_xref="GI:16877213"
/translation="MISUPPGLVLLRFLFLGLSALAPPRAQLQHLPAIRLOAVE
GGEVLPAAVTLHGVSQSPMEVPMVFFKQKEDQVLSYINGVTTSPKGVSLVY
SNPNSRLSLGLEQKSGPYSVYVQDKQKGRGHSIKTLEINLVLPAPPSCRL
QGVFVGANVTLSQSPSPKPAVQWDRQLPSQTFPPAPALDVIRGSLITNLSSM
AGVYKHAHVEGTAQCNVLEIVSTPGAAVAVAGVTVLGLGLAGLVLVHRGK
ALEPANDIKEDIAIPTLPMKPSDDTSKNSLTSVTSARALRPHGPFPAPALPT
PSLSQALPSPRLPTDTGAHPQIPSPFGVSSGSLSRMGAVPMVPAQSGSLV"
BASE COUNT 396 a 507 c 444 g 387 t
ORIGIN

Query Match 93.3%; Score 1691.8; DB 9; Length 1734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1693; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 119 gagggccatgatttcctccggggccctcgtgacaaactgctcggtttttgttct 178
Db 1 GAGGCCATGATTTCCCTCCGGGGCCCTGTTGACCAACTTGTGCGGTTTGTTCCT 60
QY 179 ggggttgatgcttcctcgccccctcgcggccccagctgcaactgcaactgcccgc 238
Db 61 GGGGCTGATGTCCTCGCGCCCCCTCGCGGGCCAGCTGCAACTGCACTGCCCGCAA 120
QY 239 ccggtgacggcggtggaggaggaggaggtggtgttcctcagcgtggtacactgacgg 298
Db 121 CCGGTTGACGGCGGTGGAGAGGGGAAAGTGTGTCTTCACGCGTGGTACACTTGCACG 180
QY 299 ggggtgtcttcctccagccagccagggaggtgcctttgtgatggtgttctcaaacag 358
Db 181 GGAGGTGTCTTCATCCAGCCATGGAGAGTCCCTTTGTGATGTGTCTTCAACAGAA 240
QY 359 agaaaggaggtacaggtgtgttcctacatcaatgggggtccaaacagcaaacctggagt 418
Db 241 AGAAAGAGGAGTACAGTGTGTCTTACATCATCAATGGGGTTCACACAGCAACCTGAGT 300
QY 419 atccttggtctactccatgcctcccggaactgtccctcgggctggaggggtctccagga 478
Db 301 ATCCTTGGTCTACTTCATGCCTCCCGGAACCTGCTCCTCGCGGTGAGGGGTCTCCAGGA 360

Db 916 TTGAGCTGGGGTGTGCTGGCTGGCTCTTTTGTATACACCGCCGGGCAAGGCCCTGG 975
 Qy 961 agagagcagcaaatatcaagagagatgccattgctcccgagaccctgcccctggccca 1020
 Db 976 AGGAGCCAGCCATGATATCAAGAGAGATGCCATTGCTCCCGAGCCCTGCTCTGGCCCA 1035
 Qy 1021 agagctcagacacaaatctcaagaatggagaccctttctctgtcacctccgacagagccc 1080
 Db 1036 AGAGCTCAGACACAAATCTCAAGAAATGGAGACCTTTCTCTGTCACTCCGCGAGAGCCC 1095
 Qy 1081 tccggccacccatggccctccagggcctgggtgattgaccccccaacccagctctctcca 1140
 Db 1096 TCCGGCCACCCCGAGCCCTCCAGGCGCTGGTGAATGACCCCGACCCAGCTCTCTCCA 1155
 Qy 1141 gccaggccctgccctcaacaagactgccacagagatggggcccaacctcaaccaatat 1200
 Db 1156 GCCAGGCCCTGCTCTCAACAAGACTGCCCAACAGAGCGGGCCCAACCTCTCAACCAATAT 1215
 Qy 1201 ccccatccctgggggtttctctcggcttgagccgcagtggtgctgacctgga 1260
 Db 1216 CTCTCATCTCTGGAGATTCTTCTCTGTGGCTTGAGCCGACAGGGTGTGGCTGTGA 1275
 Qy 1261 tgggtgctccagagtcgaagctgctctgtgatgacccacacactcattggcta 1320
 Db 1276 TGGTGGCCGCCAGAGTCAAGCTGGCTCTGTGGTGTGATGACCCGCCACTCATTTGGCTA 1335
 Qy 1321 aagatttgggggtctctctctccataaagggtcacctctcagcacagggcctgagtcag 1380
 Db 1336 AAGGANTTGGGGTCTCTCTTCTATAGGGGTCACTCTTAGCAGAGGCGCTGAGTCATG 1395
 Qy 1381 gaaag-agtcacacctcgtgaccttagtactctgcccccaacctctcttactgtggaa 1439
 Db 1396 GGAGAGTAATCACTCCAGACCTTAGTACTCTGCCCCCACTCTCTTTACTGTGGAA 1455
 Qy 1440 aaccatctcagtaagacctaaagtcacagagacagagagagagagagagagagagag 1499
 Db 1456 AACTGTCTCAGTATACCTAAGTGTCCAGAGATAGA---AGAGAGAGATGGATCTGG 1512
 Qy 1500 aattgggagagagctccaccacccctgactcctcttaagaagcagctgctgaaatta 1559
 Db 1513 AATTGGAGAGAGCTTCAACCCACCTGACTCTCTCTCTGTGAAGCAGCTGCTGAATTA 1572
 Qy 1560 gctactcaccagagtgaggggagagactccagtcactgagctcccccggccctctg 1619
 Db 1573 ACTCTCACCAGAGTGTGAGTGTGAGAGACTTCCAGTACTGAGTCTCCAGGCCCCGTTG 1632
 Qy 1620 atctgtaccacccacctatctaaacacaccccttgctcccaactccagctccctgtattga 1679
 Db 1633 ATCTGTATACCCACCTTATCTAACACACCCCTTTGCTCCCACTCCAGCTCTCTGATTGA 1692
 Qy 1680 tataacctgtcagctggcttggttaggttttactggggcagagatagggaatctctta 1739
 Db 1693 TATTAACCTGTGAGGCTGGCTTGGTTAGGTTTACTGTGGGCGAGAGATAGGAATGATTT 1752
 Qy 1740 ttaaaactaaatgaatatgtgtgttttcttcttcttcttcttcttcttcttcttcttctt 1799

Qy 1800 gttgtatgataaaa 1813
 Db 1813 GTATGTATGGATA 1826
 RESULT 8
 AX191588
 LOCUS AX191588
 15-AUG-2001
 DEFINITION Sequence 110 from Patent WO0149728.
 ACCESSION AX191588
 VERSION AX191588.1 GI:15209770
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1173)
 AUTHORS Kato, S. and Kimura, T.
 TITLE Human proteins having hydrophobic domains and dnas encoding these
 proteins
 JOURNAL Patent: WO 0149728-A 110.12-JUL-2001;
 Protegene Inc. (JP); SAGAMI CHEMICAL RESEARCH CENTER (JP)
 FEATURES
 Location/Qualifiers
 source 1..1173
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 221 a 369 c 327 g 256 t
 ORIGIN
 Query Match 64.6%; Score 1171.4; DB 6; Length 1173;
 Best Local Similarity 99.9%; Pred. No. 2.1e-261;
 Matches 1172; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 126 atgatttccctcccgggggcccttgctgacaaactgctgagggttttctctcggggctg 185
 Db 1 ATGATTTTCCCTCCCGGGGCCCTGTGTACCAACTGTGTGGGGTGTGTGTCTCTGGGGCTG 60
 Qy 186 agtgcctcgcccccctcgggccagctgcaactgcaactgcccgcacacccggttg 245
 Db 61 AGTGCCCTTCGCGCCCTCCGCGGCCCGAGCTGCAACTGTGACCTGCGCCGCCAACCGGTTG 120
 Qy 246 caggcgtgagagagaggaagtggctgtccagcgtgtacaccttgacggggagggtg 305
 Db 121 CAGGCGGTGTGAGGGGGAGGTGTGTGTCTCCAGGTGTGTACACTTTCACCGGGGAGGTG 180
 Qy 306 tcttctcccccagccatgggagggtgccccttggatgtgttcttcaaacagagaagaaag 365
 Db 181 TCTTCATCCAGGCATGGGAGGTGCCCTTTGTGTGTGTGTCTTCAACAGAGAAGAAAG 240
 Qy 366 gaggatcagggtgtgttctctacatcaatgggggtccacaataagcaaacctggagatccttg 425
 Db 241 GAGGATCAGGTGTGTCTTACATCAATGGGGGTCCACACACAGCAACCTGGGATATCCTTG 300

QY	1756	atagtgtgtt-----gttttcattggaattaaataagaatacataatgtttgtatgaaa	1811
DB	1775	ATGATGATGTTTGTGTTTATTTTGGCAATTTTCATTAAGAATACATCGCATGCAATGAA	1834
QY	1812	aa 1813	
DB	1835	AA 1836	
RESULT 10			
AP000866/c			
LOCUS	AP000866	187960 bp	DNA linear PRI
DEFINITION	Homo sapiens genomic DNA, chromosome 11q, clone:RP11-677M14, complete sequence.		
ACCESSION	AP000866		
VERSION	AP000866.4	GI:15320462	
KEYWORDS	HTG.		
SOURCE	Homo sapiens DNA, clone:RP11-677M14.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Pujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.		
TITLE	Homo sapiens genomic DNA		
JOURNAL	Published Only in Database (1999) In press		
REFERENCE	2 (bases 1 to 187960)		
AUTHORS	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Pujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-DEC-1999) Masahira Hattori, The Institute of and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan		
COMMENT	(S-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel.:81-45-503-9111, Fax:81-45-503-9170) On Aug 27, 2001 this sequence version replaced gi:9845041.		
FEATURES	Location/Qualifiers		
source	1..187960		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/chromosome="11"			
/map="11q"			
/clone="RP11-677M14"			
BASE COUNT	53649 a	43111 c	40795 g 50405 t
ORIGIN			
Query Match	45.7%	Score 828.2;	DB 9; Length 187960;
Best Local Similarity	99.6%;	Pred. No. 1.8e-181;	
Matches 830;	Conservative 0;	Mismatches 3;	Indels 0; Gaps 0;
QY	981	aaggaggatgccattgtccccggaccctgccttggcccaaggatccagacacaatctcc	1040
DB	72580	AGGAGAGATGCATTGCTCCCGGACCTGCGCTGGCGCCAGAGACTCAGACACAATCTCC	72521

ACCESSION AP000680
 VERSION AP000680.2 GI-8118868
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens DNA, clone: CMB9-25K9.
 ORGANISM Homo sapiens
 Rukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominoidea; Homo.
 1 (bases 1 to 101458)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.
 Published Only in Database (1999) In press
 2 (bases 1 to 101458)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (08-NOV-1999) Masahira Hattori, The Institute of and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Kikasso Univ., 1-15-1 Kikasso, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)
 On May 31, 2000 this sequence version replaced gi:6997534.
 ----- Genome Center
 Center: RIKEN Genomic Sciences Center (GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@sc.riken.go.jp
 ----- Project Information
 Center project name: CMB9-25K9
 Center clone name: CMB9-25K9
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 89733 bases at least Q40
 Consensus quality: 90500 bases at least Q30
 Consensus quality: 97741 bases at least Q20
 Insert size: 90056; sum-of-contigs
 Quality coverage: 4.65x in Q20 bases; sum-of-contigs

 NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and the accession number will be preserved.
 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 gaps
 sequence
 as soon as it is available and the accession number will be preserved
 1 9364 contig of 9364 bp in length
 9465 16830 contig of 7366 bp in length
 16931 23631 contig of 5701 bp in length
 23732 29864 contig of 6133 bp in length

44339 49485 contig of 100 bp
 55550 contig of 5966 bp in length
 55651 61139 contig of 5489 bp in length
 61240 65471 contig of 4232 bp in length
 65572 69042 contig of 3471 bp in length
 73512 contig of 4370 bp in length
 73613 76146 contig of 2534 bp in length
 76247 79045 contig of 2799 bp in length
 79146 82078 contig of 2933 bp in length
 82179 83583 contig of 1405 bp in length
 83684 86476 contig of 2793 bp in length
 86577 90055 contig of 2479 bp in length
 90156 92922 contig of 1767 bp in length
 92819 94573 contig of 1696 bp in length
 94674 95551 contig of 1755 bp in length
 96052 97999 contig of 1278 bp in length
 97900 100121 contig of 1748 bp in length
 100222 101458 contig of 2222 bp in length
 102022 101458 contig of 1237 bp in length
 Sequence updated (26-May-2000).
 * NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and the accession number will be preserved.
 * is not known and the accession number will be preserved.
 * is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * * * * *
 1 9364: contig of 9364 bp in length
 * 9365 9464: gap of 100 bp
 * 9465 16830: contig of 7366 bp in length
 * 16831 16930: gap of 100 bp
 * 16931 23631: contig of 5701 bp in length
 * 23632 23731: gap of 100 bp
 * 23732 29864: contig of 6133 bp in length
 * 29865 29964: gap of 100 bp
 * 29965 38284: contig of 8320 bp in length
 * 38285 38384: gap of 100 bp
 * 38385 44338: contig of 5954 bp in length
 * 44339 44438: gap of 100 bp
 * 44439 49484: contig of 5046 bp in length
 * 49485 49884: gap of 100 bp
 * 49885 55550: contig of 5966 bp in length
 * 55551 55650: gap of 100 bp
 * 55651 61139: contig of 5489 bp in length
 * 61140 61239: gap of 100 bp
 * 61240 65471: contig of 4232 bp in length
 * 65472 65571: gap of 100 bp
 * 65572 69042: contig of 3471 bp in length
 * 69043 69142: gap of 100 bp
 * 69143 73512: contig of 4370 bp in length
 * 73513 73612: gap of 100 bp
 * 73613 76146: contig of 2534 bp in length
 * 76147 76246: gap of 100 bp
 * 76247 79045: contig of 2799 bp in length
 * 79046 79145: gap of 100 bp

```

* 79146 82078: contig of 2333 bp in length
* 82079 82178: gap of 100 bp
* 82179 83583: contig of 1405 bp in length
* 83584 83683: gap of 100 bp
* 83684 86476: contig of 2793 bp in length
* 86477 86576: gap of 100 bp
* 86577 89055: contig of 2479 bp in length
* 89056 89155: gap of 100 bp
* 89156 90922: contig of 1767 bp in length
* 90923 91022: gap of 100 bp
* 91023 92718: contig of 1696 bp in length
* 92719 92818: gap of 100 bp
* 92819 94573: contig of 1755 bp in length
* 94574 94673: gap of 100 bp
* 94674 95951: contig of 1278 bp in length
* 95952 96051: gap of 100 bp
* 96052 97799: contig of 1748 bp in length
* 97800 97899: gap of 100 bp
* 97900 100121: contig of 2222 bp in length
* 100122 100221: gap of 100 bp
* 100222 101458: contig of 1237 bp in length.

FEATURES
    source             1..101458
                        .. Location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon.9606"
    /chromosomes="11"
    /map="11q24"
    /clone="CWB9-25K9"
    misc_feature       1..9364
    /note="assembly_fragment"
    misc_feature       9465..16930
    /note="assembly_fragment"
    misc_feature       16931..23631
    /note="assembly_fragment"
    misc_feature       23732..29864
    /note="assembly_fragment"
    misc_feature       29965..38284
    /note="assembly_fragment"
    misc_feature       38385..44338
    /note="assembly_fragment"
    misc_feature       44439..49484
    /note="assembly_fragment"
    misc_feature       49585..55550
    /note="assembly_fragment"
    misc_feature       55651..61139
    /note="assembly_fragment"
    misc_feature       61240..65471
    /note="assembly_fragment"
    misc_feature       65572..69042
    /note="assembly_fragment"
    misc_feature       69143..73512
    /note="assembly_fragment"
    misc_feature       73613..76146
    /note="assembly_fragment"
    misc_feature       76247..79045
    /note="assembly_fragment"
    misc_feature       79146..82078
    /note="assembly_fragment"

```

/note="interacts with HUEL gene product encoded by sequence deposited under GenBank Accession Number AF006621; similar to XPA DNA repair protein's DNA-binding domain; C4orf1"

/codon_start=1

/product="unknown"

/protein_id="AAF97236.1"

/db_xref="GI:9664853"

/translation="QEWDFPLCHLRTSPATPWPQAWCIDPHASLQPGPALTKTAH DRMGPPSTNIPHPWGPFLWLEPHGCCACGACPESSWLSGHWTPPLIG"

BASE COUNT 176 a 234 c 165 g 161 t

ORIGIN

Query Match 39.6%; Score 717.4; DB 9; Length 736;

Best Local Similarity 99.9%; Pred. No. 5.3e-156;

Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1040 caagaatggagaccccttctctgtcacctccgacagagccctccggccacccatggccc 1099

DB 1 CAAGATGGGACCCCTTCTCTGTCTCACTCCGACGAGCCCTCCGCGCCACCCCTGCGCC 60

QY 1100 tccaggcttggtagattgaccccccagccagctctctccagccagggcctgcccctacc 1159

DB 61 TCCAGGCTTGTGTGATGACCCCGCCGAGCTCTCTCAGCGAGGCGCTGCGCCCTCACC 120

QY 1160 aagactgcccacagagatggggcccccacccctcaaccaatatcccccatcccttgggggt 1219

DB 121 AAGACTGCCACGACAGATGGGCGCCACCCTCAACAAATATCCCCCATCTCTGTGGGT 180

QY 1220 ttctctctggcttgagccagatggggtgctgctgtgctgagtggtgctgcccagagctca 1279

DB 181 TTCTTCTCTGGCTTGAAGCGCATGGGTGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 240

QY 1280 agctggctctctggtatgatgaccccccacccattggtcaaggatttggggtctctcc 1339

DB 241 AGCTGGCTCTCTGT 300

QY 1340 ttctataagggtcacctctagcacagggctgagtcagtgagtgagtgagtgagtgagtgagtg 1399

DB 301 TTCTTATAGGGTTCACCTCTAGCAGAGGCGCTGTAGTCATGGGAAAGAGTCACTCTCTCTG 360

QY 1400 acccttagtacttgcggcccaactctcttactgtgggaaacacatctcagtaagacctta 1459

DB 361 ACCCTTAGTACTTGTGCCCCCACCCTCTCTTACTGTGGGAAACCACTCTCAGTAGACCTTA 420

QY 1460 agtgcaccagagacagaaggaaggaagtgagctctggaattggggaggagctccacc 1519

DB 421 AGTGTCCAGAGACAGAAAG 480

QY 1520 caccctgactctctcttatgaagccagctgctgaaatagctactctaccacaaagtgaggg 1579

DB 481 CACCCTTGACTCTCTCTATGAAGCCAGCTGTCTGAATTTAGCTACTCTCAACCAAGGTGAGG 540

QY 1580 ggcagagacttcaggtcactgagctccagggcccttgatctgtaccacccctctac 1639

DB 56972 ACCCTTAGTACTTGTGCCCCCACCCTCTCTTTACTGTGGGAAACCAATCTCAGTAGACCTTA 56913

QY 1460 agtgcaccagagacagaaggaagtgagctggaattgggagagctccacc 1519

DB 56912 ATGTCCAGAGACAG 56853

QY 1520 caccctgactctctcttatgaagccagctgctgaaatagctactcaccagagtgaggg 1579

DB 56852 CACCCTTGACTCTCTCTTATGAAGCCAGCTCTCTGAAATAGCTACTCTACCAAGAGTGAGG 56793

QY 1580 ggcagagacttcaggtcactgagctctccagggcccttgatctgtaccacccctctac 1639

DB 56792 GCGAGAGACTTCCAGTCACTGAGTCTCCAGGCGCCCTTGATCTGTATCCCAACCCCTATC 56733

QY 1640 taacacacacttggctccacacagctccctgtattgatataactgtcaggtggct 1699

DB 56732 TACACCAACCTTGGCTCCCACTCCAGCTCCCTGATTTGATATTAACCTGTCCAGGCTGGCT 56673

QY 1700 tggtaggttttactggggcagagatcaggaatctcttattataaactaaactgaatatt 1759

DB 56672 TGGTAGGTTTTACTGGGCGAG 56613

QY 1760 gtgtgttttctattgcaaatattaaataaagatacataatgtttgtatgaaaaa 1813

DB 56612 GTGTGTGTTCATTGGCAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 56559

RESULT 12

AF277292

LOCUS AF277292 736 bp mRNA linear PRI

03-AUG-2000

DEFINITION Homo sapiens C4orf1 mRNA.

ACCESSION AF277292

VERSION AF277292.1 GI:9664852

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 736)

Sim,D.L.C., Yeo,W.M. and Chow,V.T.K.

TITLE The novel human HUEL gene (C4orf1) encodes a protein that shares homology with the DNA-binding domain of the XPA DNA repair protein and displays nuclear translocation in a cell cycle-dependent manner

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 736)

Chow,V.T.K., Sim,D.L.C. and Yeo,W.M.

TITLE Direct Submission

JOURNAL Submitted (13-JUN-2000) Microbiology, National University of Singapore, 5, Science Drive 2, Singapore 117597, Singapore

FEATURES

Location/Qualifiers

1..736

1..736

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="bone marrow"

Db 541 GCGAGAGCTTCCAGTCACTGAGTCTCCAGGCCCCCTTGATCTGTATCTGATCCACCCCTATC 600

Qy 1640 taacacaccccttggctcccaatccagctccctctgattgataaactgtcaggctggct 1699
|||||

Db 601 TAAACACCCCTTGGCTCCCACTCCAGCTCCCTGTATTGATATAAACCTGTCAAGCTGGCT 660
|||||

Qy 1700 tggtagttttactggggcagagtagggaatctcttataaaactaaactgaata 1758
|||||

Db 661 TGGTTAGGTTTTACTTGGGCGAGGATAGGAATCTTTATTAATAACTTAACATGAATA 719
|||||

RESULT 13
AXI36493
LOCUS AXI36493 637 bp DNA linear PAT
30-MAY-2001
DEFINITION Sequence 415 from Patent EP1067182.
ACCESSION AXI36493
VERSION AXI36493.1 GI:1427897
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 637)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
Hayaashi,K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 415 10-JAN-2001;
Helix Research Institute (JP)
FEATURES
source 1..637
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 116 a 202 c 185 g 129 t 5 others
ORIGIN
Query Match 29.3%; Score 531.2; DB 6; Length 637;
Best Local Similarity 95.8%; Pred. No. 8.9e-113;
Matches 576; Conservative 0; Mismatches 21; Indels 4; Gaps
3;
Qy 1 ggaagccgctgggtgtcagggctcggctcccgagcagctccggcgtcgagcgt 60
Db 12 GGAGCGCGCTTGGTGTGTCAGCGGCTCGGCTCCCGGCGACGCTCGGCGCGAGCCT 71
Qy 61 cggcactgcaggtccgtcgctcccgcggtggcgccctgactccgtcccgccaggga 120
Db 72 CGGCACTGCAAGTTCGNGTCCCGCGGCTGGCGCCCTGATCTCGTCCGCGCGAGGA 131
Qy 121 gggccatgatttccctcccgggcccttgggtgaccaactgctggggttttttctctgg 180
Db 132 GGGCCATGATTTCCCTCCCGGGGCCCTCGTGTGACCACTGCTGCGGTTTTTGTCTCTGG 191
Qy 181 ggtcagtgccctcgccccccttccgcccagctgcaactgcaactgcccgaacc 240
Db 192 GGTGAGTGCCCTCGCGGCCCTCCGCGGCGCCAGCTGCACTGCACTTGGCGCCACC 251

Qy 241 ggttgcaggcgtggagggagggaagtggtgcttccagcgtgggtacacacctgcacgggg 300
|||||

Db 252 GGTTCAGCGCGGTGGAGAGAGGGGAGAGTGGTCTTCAGCGGTGGTACACCTTGACGGGG 311
|||||

Qy 301 aggtgtcttcaccccgccatggaggtgccccttggatgggttcttcaaacagaag 360
|||||

Db 312 AGGTGTCTTCATCCAGCCTGGAGGTGCCCTTTGTGATGTGGTCTTCAACACAGAAG 371
|||||

Qy 361 aaaaggaggatcagggtgtgtctctactatcaatgggggtcacaacaagcaaacctggagtat 420
|||||

Db 372 AAAAGGAGGATCAGGTGTGTCTACATCAATGGGGTCAACAAGCAAACTGGAGTAT 431
|||||

Qy 421 ccttggctactacatgcctcccggaacctgtccctgsggtggagggtctccaggaga 480
|||||

Db 432 CTTTGGTCTACTCATGCCCTCCCGAACCTGTCTCCGCTGGAGGTGCTCCAAGGAG 491
|||||

Qy 481 aagactcggccctcagctcctcg-tgaatgtcaagacadaac-aaggcaactcag 538
|||||

Db 492 AAGACTCTGGGCCCTACAGCTGCTCCGTGTGATGTGCACACCAACAGGCGAATCTTAA 551
|||||

Qy 539 gggccaca--gcatacaaaccttagaactcaatgactgggttctcctcagctccctccacc 596
|||||

Db 552 GGGCCACAGCATCAAAAACCTTAAACTCAATGTACTGTGTTCTCCACNTCTCCATCC 611
|||||

Qy 597 t 597
612 T 612

RESULT 14
AXI36640/c
LOCUS AXI36640 541 bp DNA linear PAT
30-MAY-2001
DEFINITION Sequence 562 from Patent EP1067182.
ACCESSION AXI36640
VERSION AXI36640.1 GI:14273044
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 541)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
Hayaashi,K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 562 10-JAN-2001;
Helix Research Institute (JP)
FEATURES
source 1..541
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 153 a 120 c 110 g 141 t 17 others
ORIGIN

/organism="Homo sapiens"

/db.xref="taxon:9606"

BASIS COUNT 118 a 118 c 96 g 109 t

ORIGIN

Query Match 24.3%; Score 441; DB 6; Length 441;

Best Local Similarity 100.0%; Pred. No. 7.6e-92; Mismatches 0; Indels 0; Gaps

Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps

0;

QY 1356 ctctaacacagagcctgagctcagtgaggaagagtcacactcctgaccttagtactctgc 1415

|||||

Db 1 CTCTAGCACAGAGCCCTGAGTCTATGGAAGAAGAGTACACACTCTCTGACCCCTTAGTACTCTGCG 60

|||||

QY 1416 cccacactctcttactgtgggaagaaacatctcagtaagacactaaagtgtccagagacag 1475

|||||

Db 61 CCCACCTCTCTTACTGTGGGAAACCATCTCAGTAAGACCTAAGTGTCCAGAGACAG 120

|||||

QY 1476 aaggagaagagagtgagctcggaattggaggagcctccacccacccctgacctctcc 1535

|||||

Db 121 AAGCAGAGAGGAGTGGATCTGGAATVGGAGGAGCTCCACCCACCCCTGACTCTCTCC 180

|||||

QY 1536 ttatgaagccagctgctgnaattgactcaccagagtgaggagggcagagactccagt 1595

|||||

Db 181 TTATGAAGCCAGCTGCTGAAATTAGCTACTCACCAGAGTGAGGGGCGAGAGACTTCCAGT 240

|||||

QY 1596 cactgagctccccagggcccccttgatctgacccacccctatctaaccacccctggc 1655

|||||

Db 241 CACTGAGTCTCCAGGCCCCCTTGATCTGTACCCACCCCTATCTAACACACCCCTTGCG 300

|||||

QY 1656 tccactccagctccctgtatgatataaacctgctcaggctggctgggttaggtttactg 1715

|||||

Db 301 TCCACTCCAGCTCCCTGTATGATATTAACCTGTGCGCTGGCTGGCTGGTGGTGGTGGT 360

|||||

QY 1716 gggcagagagatagggaatctcttattataaactaaactgaaatgctgtgtttcattg 1775

|||||

Db 361 GGGCAGAGGATAGGGAATCTCTATTAATAAATAACATGAATAATGTGTGTGTTCATTG 420

|||||

QY 1776 caaatattaaataagatacat 1796

|||||

Db 421 CAATTTTAATAAAGATACAT 441

|||||

Search completed: August 19, 2002, 16:13:11

Job time: 4324 sec

Db 534 TTAAAGNGGTTTTTGGNANGAGNACCCCAACCATNATTTGGTTAAAGGA-TTGGGGTTTT 476

|||||

QY 1337 tctctctataagggctacctctagcacagagcctgagctcatgggaagagctcacactc 1396

|||||

Db 475 TCTTTCTATTAAGGCTCATTTTATGACACAGAGGCTGAGTCTATGGAAGAGTCAATTC 416

|||||

QY 1397 ctgaccttagtactctgccccactctcttactgtgggaagaaacactcagtaagac 1456

|||||

Db 415 NTGACCTTAGTATTNTGCCCCACTTTTTTTTANTGTGGGAAACCATNTCAGTAAGAC 356

|||||

QY 1457 ctaagtgt-cagagagacagagagaggaagagtgagctcggaattggaggagcctc 1515

|||||

Db 355 NTAAGTGTCTCCAGGACACAGAGAGAGGAGTGTGATNTGGAATVGGAGGAGTTC 296

|||||

QY 1516 caccacccctgactcctcttataagccagcgtctgnaattgactactcaccagagt 1575

|||||

Db 295 CACCCACCCCTGATCTCTTATGAAGCCAGCTGCTGAAATTAGCTANTCACCAGAGT 236

|||||

QY 1576 gaggggcagagactccagctcactgagctctccaggcccccttgatctgtaccacccc 1635

|||||

Db 235 GAGGGGACAGAGCTTCAGTCTAGTGTTCACAGGCCCCCTTGATNTGTATCCACCC 176

|||||

QY 1636 tatctaaccacacccttggtccctccactccagctccctgtattgatataaacctgtcaggt 1695

|||||

Db 175 TATNTAACACACCTTGNTCCACTCCAGTCCCTGATNTGATATTAACCTGTCCAGCT 116

|||||

QY 1696 ggcctggtaggttttactggggcagagagatagggaatctcttataaactaaatgaa 1755

|||||

Db 115 GCGTGTGTTAGGTTTTTACTGGGGCAGAGGATAGGGAATTTTTTATTAATAACTAACATGAA 56

|||||

QY 1756 atagtgtttttcatttgcaatttaataaagatacataatgtttgtatga 1809

|||||

Db 55 ATATGTGTGTTTTCATTTGCAAAATTTAAATAAAGATACATATGTGTGTATGA 2

|||||

RESULT 15

AX332845

LOCUS AX332845 441 bp DNA linear PAT

09-JAN-2002

DEFINITION Sequence 3354 from Patent WO0194629.

ACCESSION AX332845

VERSION AX332845.1 GI:18123479

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,

Horrigian, S., Soppet, D.R. and Weaver, Z.

TITLE

Cancer gene determination and therapeutic screening using

signature

gene sets

JOURNAL Patent: WO 0194629-A 3354 13-DEC-2001;

OM nucleic - nucleic search, using sw model

Run on: August 19, 2002, 15:04:12 ; Search time 220.41 Seconds
(without alignments)
14122.616 Million cell updates/sec

Title: US-09-902-759-38

Perfect score: 1813
Sequence: 1 ggaagccgcctgggtcgag.....cataatgtttgtatgaaaa 1813

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_032802.*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Score Match Length DB ID Description

Result No.	Score	Match	Length	DB	ID	Description
1	1813	100.0	1813	20	AAK52221	Protein PRO246 cDN
2	1813	100.0	1813	20	AAK28436	EGF-like homologue
3	1813	100.0	1813	21	AAA30052	Human PRO246 nucle
4	1813	100.0	1813	22	AAS21412	Human cDNA sequenc
5	1813	100.0	1813	22	AAK60372	PRO246 coding sequ
6	1813	100.0	1813	22	AAC67040	Nucleotide sequenc
7	1813	100.0	1813	22	AAK72379	Human PRO246 cDNA.
8	1813	100.0	1813	22	AAC97441	Human angiogenesis
9	1809	99.8	1821	22	AAK93785	Human cDNA encodin
10	1806.6	99.6	1827	22	AAH02949	Human shear stress
11	1804.8	99.5	1816	22	AAD12605	Human protein havi
12	1802	99.4	1954	21	AAA23441	cDNA encoding huma
13	1783.8	98.4	1932	21	AAC26278	Human secreted pro
14	1760.4	97.1	1831	22	AAC85076	Atherosclerosis-as
15	1757.6	96.9	1869	22	AAK44978	Human INTERCEPT 25
16	1756	96.9	1869	22	AAK45014	Human secreted pro
17	1756	96.9	1869	22	AAK45015	Human secreted pro
18	1756	96.9	1869	22	AAK45016	Human secreted pro
19	1756	96.9	1869	22	AAK45017	Human secreted pro
20	1440.8	79.5	1748	22	ABA09181	Human viral recept
21	1440.8	79.5	1748	22	AAI59707	Human polynucleoti
22	1376	75.9	1387	20	AAK57000	Human viral recept
23	1083.2	59.7	1290	20	AAZ00447	Human secreted pro
24	1069.8	59.0	1110	22	AAK44979	Human INTERCEPT 25
25	1068.2	58.9	1110	22	AAK45046	Human secreted pro
26	1068.2	58.9	1110	22	AAK45047	Human secreted pro
27	1068.2	58.9	1110	22	AAK45048	Human secreted pro
28	1068.2	58.9	1110	22	AAK45049	Human secreted pro
29	1010.4	55.7	1606	22	AAI57921	Human polynucleoti
30	863.2	47.6	1846	22	AAK45020	Murine secreted pr
31	861.6	47.5	1846	22	AAK44981	Murine INTERCEPT 2
32	861.6	47.5	1846	22	AAK45018	Murine secreted pr
33	861.6	47.5	1846	22	AAK45021	Murine secreted pr
34	860	47.4	1846	22	AAK45019	Murine secreted pr
35	730.2	40.3	1182	22	AAK45052	Murine secreted pr
36	728.6	40.2	1182	22	AAK44982	Murine INTERCEPT 2
37	728.6	40.2	1182	22	AAK45050	Murine secreted pr
38	728.6	40.2	1182	22	AAK45053	Murine secreted pr
39	727	40.1	1182	22	AAK45051	Murine secreted pr
40	578.4	30.9	1288	22	AAD10122	Mouse 10.3 kDa pro
41	531.2	29.3	637	22	AAK93981	Primer specific fo
42	453	25.0	541	22	AAK94128	Primer specific fo
43	257.2	14.2	571	22	AAK97944	Murine 7-transmemb
44	226.4	12.5	564	22	AAH97943	Murine 7-transmemb
45	208	11.5	533	22	AAH97945	Murine 7-transmemb

ALIGNMENTS

RESULT 1
AAK52221
ID AAK52221 standard; DNA; 1813 BP.

DT 25-JUN-1999 (first entry)
 DE Protein PRO246 cDNA clone DNA35639-1172.
 XX
 XX Secreted protein; transmembrane protein; human; enterocolitis;
 KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
 KW congenital microvillus atrophy; skin disease; cell growth;
 KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
 KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
 KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
 KW anti-thrombotic; wound healing; tissue repair; ss.
 XX
 OS Homo sapiens.
 XX
 XX W09914328-A2.
 XX
 PD 25-MAR-1999.
 XX
 XX 16-SEP-1998; ..98MO-US19330.
 XX
 XX 25-NOV-1997; 97US-0066840.
 PR 17-SEP-1997; 97US-0059113.
 PR 17-SEP-1997; 97US-0059115.
 PR 17-SEP-1997; 97US-0059117.
 PR 17-SEP-1997; 97US-0059119.
 PR 17-SEP-1997; 97US-0059121.
 PR 17-SEP-1997; 97US-0059122.
 PR 17-SEP-1997; 97US-0059184.
 PR 18-SEP-1997; 97US-0059263.
 PR 18-SEP-1997; 97US-0059266.
 PR 15-OCT-1997; 97US-0062125.
 PR 17-OCT-1997; 97US-0062285.
 PR 17-OCT-1997; 97US-0062287.
 PR 21-OCT-1997; 97US-0063486.
 PR 24-OCT-1997; 97US-0062814.
 PR 24-OCT-1997; 97US-0062816.
 PR 24-OCT-1997; 97US-0063045.
 PR 24-OCT-1997; 97US-0063120.
 PR 24-OCT-1997; 97US-0063121.
 PR 24-OCT-1997; 97US-0063127.
 PR 24-OCT-1997; 97US-0063128.
 PR 27-OCT-1997; 97US-0063329.
 PR 27-OCT-1997; 97US-0063327.
 PR 28-OCT-1997; 97US-0063541.
 PR 28-OCT-1997; 97US-0063542.
 PR 28-OCT-1997; 97US-0063544.
 PR 28-OCT-1997; 97US-0063549.
 PR 28-OCT-1997; 97US-0063550.
 PR 28-OCT-1997; 97US-0063564.
 PR 29-OCT-1997; 97US-0063435.
 PR 29-OCT-1997; 97US-0063704.
 PR 29-OCT-1997; 97US-0063732.
 PR 29-OCT-1997; 97US-0063738.
 PR 29-OCT-1997; 97US-0063734.
 PR 29-OCT-1997; 97US-0064215.

03-NOV-1997; 97US-0064248.
 07-NOV-1997; 97US-0064809.
 12-NOV-1997; 97US-0065186.
 17-NOV-1997; 97US-0065846.
 18-NOV-1997; 97US-0065693.
 21-NOV-1997; 97US-0066120.
 21-NOV-1997; 97US-0066364.
 24-NOV-1997; 97US-0066772.
 24-NOV-1997; 97US-0066466.
 24-NOV-1997; 97US-0066770.
 24-NOV-1997; 97US-0066511.
 24-NOV-1997; 97US-0066453.

(GETH) GENENTECH INC.
 Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
 WPI; 1999-229533/19.
 P-PSDB; AAV13351.
 New isolated human genes and polypeptides used in, e.g. treatment of
 gastrointestinal ulceration
 Claim 2; Fig 16; 320pp; English.
 AAX52213-74 encode secreted and transmembrane human proteins, and are
 obtained from cDNA libraries, prepared from fetal lung, fetal kidney,
 fetal brain, fetal liver and fetal retina. The encoded polypeptides
 have specific uses based on their homology to known polypeptides,
 e.g. PRO211 and PRO217 can be used for disorders associated with the
 preservation and maintenance of gastrointestinal mucosa and the repair
 of acute and chronic mucosal lesions (e.g. enterocolitis,
 Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
 microvillus atrophy), skin diseases associated with abnormal
 keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as
 lung squamous cell carcinoma of the vulva and gliomas), potent effects
 cell growth and development, diseases related to growth or survival of
 nerve cells including Parkinson's disease, Alzheimer's disease, ALS,
 neuropathies or cancer. PRO265 can be used as for fibromodulin, e.g. for
 reducing dermal scarring. PRO264 can be used as a target for anti-tumor
 drugs. PRO533 may be used in the treatment of Usher Syndrome or Atrophia
 areata; PRO269 can be used as an anti-thrombotic agent; PRO287
 polypeptides and portions may have therapeutic applications in wound
 healing and tissue repair; PRO317 can be used for treating problems of
 the kidney, uterus, endometrium, blood vessels, or related tissue, e.g.
 in the heart of genital tract.

Sequence 1813 BP; 368 A; 559 C; 484 G; 402 T; 0 other;
 Query Match 100.0%; Score 1813; DB 20; Length 1813;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 1813; Conservative 0; Mismatches 0; Indels 0; Gaps
 0;

QY 1 ggagccgcccctgggtgtcaaggcctcgagctccgcgcgcagcctccggccgctgcgcgcagcct 60
 Db 1 ggagccgcccctgggtgtcagcggctcggtccgcgcgcagcctccggccgctgcgcgcagcct 60
 QY 61 cggcaacctgcaggtccgtgcgtccgcgcgcagcctcgcccccctgactccgtccgcgcagagga 120
 Db 61 cggcaacctgcaggtccgtgcgtccgcgcgcagcctcgcccccctgactccgtccgcgcagagga 120
 QY 121 gggccatgattccctccggggccctcgatgaccaactgctgcgggtttttgttccctgg 180
 Db 121 gggccatgattccctccggggccctcgatgaccaactgctgcgggtttttgttccctgg 180
 QY 181 ggctgagtcctcgcccccctccggcccgagctgcaactgcaactgcccgcaccc 240
 Db 181 ggctgagtcctcgcccccctccggcccgagctgcaactgcaactgcccgcaccc 240
 QY 241 ggtgcaggcgggtggaagggaaggagtggtgcttccagcgtggtacaccttgcacgggg 300
 Db 241 ggtgcaggcgggtggaagggaaggagtggtgcttccagcgtggtacaccttgcacgggg 300
 QY 301 aggtgtctatccagccatggaggtgccttgcgtgattgggttttcaaacagaaag 360
 Db 301 aggtgtctatccagccatggaggtgccttgcgtgattgggttttcaaacagaaag 360
 QY 361 aaaggagagatcaggtgtctctcatcaatggggtcacaaagcaaacctggagat 420
 Db 361 aaaggagagatcaggtgtctctcatcaatggggtcacaaagcaaacctggagat 420
 QY 421 ccttggtctactcctcccgaaacctgcccgcggctccctgcggctggagggtctccagagga 480
 Db 421 ccttggtctactcctcccgaaacctgcccgcggctccctgcggctggagggtctccagagga 480
 QY 481 aagactctggcccctacagctgtccgtgaatgtgcaagacacaaaggcaaatctagg 540
 Db 481 aagactctggcccctacagctgtccgtgaatgtgcaagacacaaaggcaaatctagg 540
 QY 541 gccacagatcaaaacctagaaactcaatgactggttcctccagctctccatcctgcc 600
 Db 541 gccacagatcaaaacctagaaactcaatgactggttcctccagctctccatcctgcc 600
 QY 601 gtctccaggtgtgcccatgtgggggcaaacgtgacctgaggtgccagttcccaagga 660
 Db 601 gtctccaggtgtgcccatgtgggggcaaacgtgacctgaggtgccagttcccaagga 660
 QY 661 gtaagcccgctgtccaataccagtggaatcgaggttccatctccagactttttctg 720
 Db 661 gtaagcccgctgtccaataccagtggaatcgaggttccatctccagactttttctg 720
 QY 721 caccagcttagatgcatccgtgggtctttaagcctcaccacactttctgtcttcattgg 780
 Db 721 caccagcttagatgcatccgtgggtctttaagcctcaccacactttctgtcttcattgg 780
 QY 781 ctggagcttatgtctgaaggcccaaatgaggtgggcaactgcccgaatgtaagtgaac 840
 Db 781 ctggagcttatgtctgaaggcccaaatgaggtgggcaactgcccgaatgtaagtgaac 840
 QY 841 tggaaagtgaacaggccctgagctgcagtgattgctggagctgtgtgggtaccctgg 900

Db 601 gtccaccaggggtgcccacatgctgggggcaaacgtagccctgagctgcacgtcccaaggga 660
 Qy 661 gtaagcccgctgtccaatacagtgaggtgcgagcttcacatccctccagactttcttg 720
 Db 661 gtaagcccgctgtccaatacagtgaggtgcgagcttcacatccctccagactttcttg 720
 Qy 721 caccagcattagatgtcatccgtgggtctcttaagctccaccaaccttgcctttccatgg 780
 Db 721 caccagcattagatgtcatccgtgggtctcttaagctccaccaaccttgcctttccatgg 780
 Qy 781 ctggagctctatgtctgaagcccccacaatgaaggtgggcactgcccacatgtaatgtagcgc 840
 Db 781 ctggagctctatgtctgaagcccccacaatgaaggtgggcactgcccacatgtaatgtagcgc 840
 Qy 841 tggaggtgagcacagccgtcgagctgcagtgaggtggtctgctggagctgttctgggtaccctgg 900
 Db 841 tggaggtgagcacagccgtcgagctgcagtgaggtggtctgctggagctgttctgggtaccctgg 900
 Qy 901 ttgagctgggggtgtgctgggtgggtcctctgtaccacgcgggggcaagggccctgg 960
 Db 901 ttgagctgggggtgtgctgggtgggtcctctgtaccacgcgggggcaagggccctgg 960
 Qy 961 agggagcagccaatgatatacaggagatgccattgtccccggaccctgccctggccca 1020
 Db 961 agggagcagccaatgatatacaggagatgccattgtccccggaccctgccctggccca 1020
 Qy 1021 agagctcagacacaaatctcaagaatgggacccttctctgtcactctcgcagagccc 1080
 Db 1021 agagctcagacacaaatctcaagaatgggacccttctctgtcactctcgcagagccc 1080
 Qy 1081 tccggccacccatggcctcccaagctggtgcattgaaccccaagccagctctccca 1140
 Db 1081 tccggccacccatggcctcccaagctggtgcattgaaccccaagccagctctccca 1140
 Qy 1141 gccaggccctgcctcaccagactgccacagacagatggggccaccctccaaccaat 1200
 Db 1141 gccaggccctgcctcaccagactgccacagacagatggggccaccctccaaccaat 1200
 Qy 1201 ccccatccctgggtgggtttctctctgggttgagcccgatgggtgctgctctgta 1260
 Db 1201 ccccatccctgggtgggtttctctctgggttgagcccgatgggtgctgctctgta 1260
 Qy 1261 tgggtgctgccagagtcgaagctgctctctgggtgatgaccccaacctatggcta 1320
 Db 1261 tgggtgctgccagagtcgaagctgctctctgggtgatgaccccaacctatggcta 1320
 Qy 1321 aaggatttgggtgtctctctcctaagggtcaactcttagcacagagcctgagtcag 1380
 Db 1321 aaggatttgggtgtctctctcctaagggtcaactcttagcacagagcctgagtcag 1380
 Qy 1381 ggaagagtcacactctgacccttagtactctgccccaccctcttctacttgaggaa 1440
 Db 1381 ggaagagtcacactctgacccttagtactctgccccaccctcttctacttgaggaa 1440
 Qy 1441 accatctcagtaagacctaaggtgtccaggagacagaaggaaggaagtgatctgga 1500
 Db 1441 accatctcagtaagacctaaggtgtccaggagacagaaggaaggaagtgatctgga 1500

PT diagnosis and treatment of neoplastic cell growth and proliferation -
 XX
 PS Example 8; Fig 15; 200pp; English.
 XX
 CC This sequence represents a human PRO246 nucleotide sequence. PRO246 is
 CC probably a cell surface virus receptor. The invention relates to
 CC isolated
 CC antibodies which bind to a polypeptide. The "pro" polypeptides are
 CC encoded by genes which are over expressed in the genome of tumour cells.
 CC Vectors and host cells comprising the nucleic acid encoding the
 CC antibodies are used in the production of the antibodies. The antibodies
 CC and nucleic acids encoding them are used for diagnosing a tumour in a
 CC mammal. The antibodies are used for inhibiting the growth of tumour
 CC cells
 CC and identifying compounds that inhibit a biological or immunological
 CC activity of and/or expression of a PRO187, PRO533, PRO214, PRO240,
 CC PRO211, PRO230, PRO261, PRO246 or PRO317 polypeptide. The antibody can
 CC be
 CC used in antibody dependent enzyme mediated prodrug therapy (ADEPT) by
 CC conjugating the antibody to a prodrug-activating enzyme which converts
 CC a
 CC prodrug to an anti-cancer drug. The antibodies can be fluorescently
 CC labelled and monitored by light microscopy, flow cytometry or
 CC fluorimetry
 CC for diagnosis and prognosis of tumours.
 XX
 SQ Sequence 1813 BP; 368 A; 559 C; 484 G; 402 T; 0 other;

Query Match 100.0%; Score 1813; DB 21; Length 1813;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1813; Conservative 0; Mismatches 0; Indels 0; Gaps
 0;
 Qy 1 ggaacccctgggtgcagcggctcggtcccgccagcgtccggtcgagcgcct 60
 Db 1 ggaacccctgggtgcagcggctcggtcccgccagcgtccggtcgagcgcct 60
 Qy 61 cggcacctcaggtccggtcccgccggtcggtcccgccctgactccgcccagggga 120
 Db 61 cggcacctcaggtccggtcccgccggtcggtcccgccctgactccgcccagggga 120
 Qy 121 gggccatgattccctcccggggcccctggcgacaaactgctgagttttgtcctgg 180
 Db 121 gggccatgattccctcccggggcccctggcgacaaactgctgagttttgtcctgg 180
 Qy 181 ggcctgactgcctcgcccccctcgsgcccaactgcgaactgcaattcccgcaacc 240
 Db 181 ggcctgactgcctcgcccccctcgsgcccaactgcgaactgcaattcccgcaacc 240
 Qy 241 ggttgacggcggtgagggagggaagtgggtcttcacagcgtggtacaccttgacgggg 300
 Db 241 ggttgacggcggtgagggagggaagtgggtcttcacagcgtggtacaccttgacgggg 300
 Qy 301 aggtgtcttcaccccgccatgggaggtgcctttgtgagtggtttctcaaacagaaag 360
 Db 301 aggtgtcttcaccccgccatgggaggtgcctttgtgagtggtttctcaaacagaaag 360

Qy 361 aaaaggaggtcaggtgtgtgtccacatacaatggggtcacaacagcaaaactcggagtat 420
 Db 361 aaaaggaggtcaggtgtgtgtccacatacaatggggtcacaacagcaaaactcggagtat 420
 Qy 421 ctttggtctactccatgcctcccggaacactgctccctgcggtcgaggggtcccgaggaga 480
 Db 421 ctttggtctactccatgcctcccggaacactgctccctgcggtcgaggggtcccgaggaga 480
 Qy 481 aagactcggccctcactagctgctcogtgaatgcaagaacaacaggcaaatcctagg 540
 Db 481 aagactcggccctcactagctgctcogtgaatgcaagaacaacaggcaaatcctagg 540
 Qy 541 gccacagctcaaaactctagaactcaatgactcgtgtgtcctccagctccctccatcctgcc 600
 Db 541 gccacagctcaaaactctagaactcaatgactcgtgtgtcctccagctccctccatcctgcc 600
 Qy 601 gtctccaggtgtcctccatgctgggggcaaacgctgacctgacctgacctcccaaggga 660
 Db 601 gtctccaggtgtcctccatgctgggggcaaacgctgacctgacctgacctcccaaggga 660
 Qy 661 gtaagccctcgtccataaccagctgggtcggcagctccatctccagactctctttg 720
 Db 661 gtaagccctcgtccataaccagctgggtcggcagctccatctccagactctctttg 720
 Qy 721 caccagcattagatgctacccgtgggtctttaagcctcccaacacttctcctccatcgg 780
 Db 721 caccagcattagatgctacccgtgggtctttaagcctcccaacacttctcctccatcgg 780
 Qy 781 ctggagctctatgtctgcaaggcccaatgaggtggcactgcccaatgtaagtgagcgc 840
 Db 781 ctggagctctatgtctgcaaggcccaatgaggtggcactgcccaatgtaagtgagcgc 840
 Qy 841 tggagtgagcacagggcctggagctgactggtgtgctgagagctgtgtgggtaccctgg 900
 Db 841 tggagtgagcacagggcctggagctgactggtgtgctgagagctgtgtgggtaccctgg 900
 Qy 901 ttggactgggggtgctgggtgggtcctctctgtaccaccccggggcaaggccctgg 960
 Db 901 ttggactgggggtgctgggtgggtcctctctgtaccaccccggggcaaggccctgg 960
 Qy 961 agggagcagccaatgatcaaggagtagcctgctcccggaacctgacctgggcccga 1020
 Db 961 agggagcagccaatgatcaaggagtagcctgctcccggaacctgacctgggcccga 1020
 Qy 1021 agagctcagacacaaatctccaaagaatgggaccccttctctgtcactccgacagcgc 1080
 Db 1021 agagctcagacacaaatctccaaagaatgggaccccttctctgtcactccgacagcgc 1080
 Qy 1081 tcgggccaaccaatggcctcccgagcctgggtgactgacccccacgcccagctctcca 1140
 Db 1081 tcgggccaaccaatggcctcccgagcctgggtgactgacccccacgcccagctctcca 1140
 Qy 1141 gccagggcctgcctcaccagactgccacagacagatggggccaccctcaaccaaat 1200
 Db 1141 gccagggcctgcctcaccagactgccacagacagatggggccaccctcaaccaaat 1200

muscle;

KW adipocyte; A-peptide; factor VIIA; gene therapy; aa.

XX Homo sapiens.

OS WO200140466-A2.

PX

PD 07-JUN-2001.

PF 01-DEC-2000; 200OWO-US32678.

XX 01-DEC-1999; 99WO-US28301.

PR 01-DEC-1999; 99WO-US28634.

PR 02-DEC-1999; 99WO-US28551.

PR 02-DEC-1999; 99WO-US28564.

PR 02-DEC-1999; 99WO-US28565.

PR 02-DEC-1999; 99US-0170262.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 20-DEC-1999; 99WO-US30999.

PR 30-DEC-1999; 99WO-US31243.

PR 06-JAN-2000; 200OWO-US00277.

PR 06-JAN-2000; 200OWO-US00376.

PR 11-FEB-2000; 200OWO-US03565.

PR 18-FEB-2000; 200OWO-US04341.

PR 18-FEB-2000; 200OWO-US04342.

PR 22-FEB-2000; 200OWO-US04414.

PR 24-FEB-2000; 200OWO-US04914.

PR 24-FEB-2000; 200OWO-US05004.

PR 01-MAR-2000; 200OWO-US05601.

PR 20-MAR-2000; 200OWO-US07377.

PR 21-MAR-2000; 200OWO-US07532.

PR 30-MAR-2000; 200OWO-US08439.

PR 17-MAY-2000; 200OWO-US13705.

PR 22-MAY-2000; 200OWO-US14042.

PR 30-MAY-2000; 200OWO-US14941.

PR 02-JUN-2000; 200OWO-US15264.

PR 10-NOV-2000; 200OWO-US30873.

XX

PA (GETH) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

XX
DR WFI; 2001-408281/43.
DR P-FSDB; RAU12340.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect other PRO polypeptides, link bioactive molecules to cells expressing PRO polypeptides, and detect the presence of mammalian tumours e.g. lung, breast, prostate, cervical

XX Claim 3; Fig 337; 813pp; English.

PS
CC AA521244-AA521518 encode for novel human secretory and transmembrane PRO polypeptides. The PRO polypeptides are useful to detect other CC PRO polypeptides, to link bioactive molecules to cells expressing CC PRO polypeptides, ear proliferation; glucose; free fatty acid; skeletal cartilage; ear:

QY 1201 ccccccacctggcgagggtttcttcctcctgagccacatgggcgtgactgtga 1260
Db 1201 cccccacctggcgagggtttcttcctcctgagccacatgggcgtgactgtga 1260
QY 1261 tggtagctgcccaagtcgaagtggcttcttgattgatgaccacacacatcatggcta 1320
Db 1261 tggtagctgcccaagtcgaagtggcttcttgattgatgaccacacacatcatggcta 1320
QY 1321 agsgatttggggtctctcttctataagggtcacctctgacacagaggtgagtcag 1380
Db 1321 agsgatttggggtctctcttctataagggtcacctctgacacagaggtgagtcag 1380
QY 1381 ggaagaagtcacaactctgaccttagtaactctgcccaccctctcttacttgtggaaa 1440
Db 1381 ggaagaagtcacaactctgaccttagtaactctgcccaccctctcttacttgtggaaa 1440
QY 1441 accatctcagtaagaactaaagtgtccaggagacagaggaaggaggtgacttggga 1500
Db 1441 accatctcagtaagaactaaagtgtccaggagacagaggaaggaggtgacttggga 1500
QY 1501 attggaggagctccaccacacctgaactctctcttatgaagccagctgtgaaattag 1560
Db 1501 attggaggagctccaccacacctgaactctctcttatgaagccagctgtgaaattag 1560
QY 1561 ctactccaagaagtgggggagagacttccagtcactagctctccagggcccccttga 1620
Db 1561 ctactccaagaagtgggggagagacttccagtcactagctctccagggcccccttga 1620
QY 1621 ttctgtacccaccctctatccaccaccccttggctcccatcagctccctgtattgat 1680
Db 1621 ttctgtacccaccctctatccaccaccccttggctcccatcagctccctgtattgat 1680
QY 1681 ataacctgtcaggctggctgttaggttttactggggcagagatataggaaactcttat 1740
Db 1681 ataacctgtcaggctggctgttaggttttactggggcagagatataggaaactcttat 1740
QY 1741 taaactaacatgaataatgtgtgttttcattgtccaatttaaataaagatacataatg 1800
Db 1741 taaactaacatgaataatgtgtgttttcattgtccaatttaaataaagatacataatg 1800
QY 1801 ttgtatgaaaaa 1813
Db 1801 ttgtatgaaaaa 1813

RESULT 4.
ID AA521412
AC AA521412 standard; CDNA; 1813 BP.
XX
AC AA521412;
XX
DT 24-OCT-2001 (first entry)
DE Human CDNA sequence encoding for PRO246 polypeptide.
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal

PRO polypeptides, and to detect the presence of mammalian lung, colon, breast, prostate, rectal, cervical or liver tumours by comparing PRO polypeptide expression in a cell sample to that in a control sample. Some of the 275 sequences are also useful to stimulate the release of tumour necrosis factor- α (TNF- α) from human blood, the proliferation or differentiation of chondrocytes, the proliferation or gene expression in pericyte cells, the release of proteoglycans from cartilage, the proliferation of inner ear utricular supporting cells or T-lymphocytes, the release of a cytokine from peripheral blood monocytes (PMNCs), or the proliferation of endothelial cells. Some of the PRO polypeptides may modulate glucose or free fatty acid uptake by skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide to factor VIIA. The PRO polypeptides can be used in assays to identify molecules involved in binding interactions. The polynucleotides encoding PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.

Sequence 1813 BP; 368 A; 559 C; 484 G; 402 T; 0 other:

```

Query Match      100.0%; Score 1813; DB 22; Length 1813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1813; Conservative 0; Mismatches 0; Indels 0; Gaps
0;

QY 1 ggaagcgcctgggtgtcagcgagtcgggtccgcgcgacgctccgcgcgtcgcgagcgt 60
Db 1 ggaagcgcctgggtgtcagcgagtcgggtccgcgcgacgctccgcgcgtcgcgagcgt 60

QY 61 cggacctggaagtccggtgcctccgcggctgcgcacctgactccgtccgcggccaggga 120
Db 61 cggacctggaagtccggtgcctccgcggctgcgcacctgactccgtccgcggccaggga 120

QY 121 gggccatgatttccctccggggccctcggcgaccacctgctgcggttttgtctcgg 180
Db 121 gggccatgatttccctccggggccctcggcgaccacctgctgcggttttgtctcgg 180

QY 181 ggcctgagtgccctcgcgcgccctccgcgggccagctgcaactgcacttgcgcgcaacc 240
Db 181 ggcctgagtgccctcgcgcgccctccgcgggccagctgcaactgcacttgcgcgcaacc 240

QY 241 ggctgcagcggcggtagggaggagggaagtgggtcctccagctggtaaccttgcacgggg 300
Db 241 ggctgcagcggcggtagggaggagggaagtgggtcctccagctggtaaccttgcacgggg 300

QY 301 aggtgtcttcatccagcaatgggagggtgcctcttgatgtggtttcttcaacagagaag 360
Db 301 aggtgtcttcatccagcaatgggagggtgcctcttgatgtggtttcttcaacagagaag 360

QY 361 aaaaggaggatcagggtgtgtctcatcaatggggttcacacgaagcaaacctggagat 420
Db 361 aaaaggaggatcagggtgtgtctcatcaatggggttcacacgaagcaaacctggagat 420

QY 421 ccttggtctactcatgcctcccggaacctgtccctgcgctggagggtctccaggaga 480
Db 421 ccttggtctactcatgcctcccggaacctgtccctgcgctggagggtctccaggaga 480

```


13-SEP-1999; 99WO-US20944.
 15-SEP-1999; 99WO-US21090.
 05-OCT-1999; 99WO-US23089.
 29-NOV-1999; 99WO-US28214.
 30-NOV-1999; 99WO-US28313.
 02-DEC-1999; 99WO-US28564.
 (GETH) GENENTECH INC.
 Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;
 WPI; 2001-091968/10.
 P-PSDB; AAB68599.
 New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,
 useful for diagnosing and treating cancers -
 Claim 50; Fig 15; 196pp; English.
 The present invention relates to PRO proteins and coding sequences. The
 present sequence is the coding sequence for one such PRO protein.
 It was found that the PRO genes are amplified in the genome of tumour
 cells. The gene amplification is expected to be associated with the
 overexpression of the gene product and contributes to tumorigenesis.
 Therefore, antagonists of PRO proteins are useful for the treatment of
 benign or malignant tumours, leukemias, lymphoid malignancies and other
 disorders such as neuronal, glial, astrocytic, hypothalamic, glandular,
 epithelial, inflammatory and immunologic disorders.
 Sequence 1813 BP; 368 A; 559 C; 484 G; 402 T; 0 other;

1321 aaggatttgggggtctctcttctctataaagggtacacctcttagcacagggccgtcgagctatg 1380
 1381 ggaagagtcacactctgaccccttagtactctgcccacacctctcttactgtgggaaa 1440
 1381 ggaagagtcacactctgaccccttagtactctgcccacacctctcttactgtgggaaa 1440
 1441 accatctcagtaagaactgaagtcctcagagacagaagagaaggaagtggatctcggg 1500
 1441 accatctcagtaagaactgaagtcctcagagacagaagagaaggaagtggatctcggg 1500
 1501 attggaggagcctccacccacctgactctctctatgaagccagctgctggaataatg 1560
 1501 attggaggagcctccacccacctgactctctctatgaagccagctgctggaataatg 1560
 1561 ctactcacaagactgaagggcagagactccagtcactgagctctccagggcccttga 1620
 1561 ctactcacaagactgaagggcagagactccagtcactgagctctccagggcccttga 1620
 1621 tctgtacccacccctctatcaacacaccccttggtctccactccagctccctgtatgat 1680
 1621 tctgtacccacccctctatcaacacaccccttggtctccactccagctccctgtatgat 1680
 1681 ataactgtcaggtgcttgtagtttactggggcagagatagggaatctcttat 1740
 1681 ataactgtcaggtgcttgtagtttactggggcagagatagggaatctcttat 1740
 1741 taaactaacatgaatgctgtgtttctatgtgcaaatatttaataaatacataatg 1800
 1741 taaactaacatgaatgctgtgtttctatgtgcaaatatttaataaatacataatg 1800
 1801 tttgtatgaaaaa 1813
 1801 tttgtatgaaaaa 1813
 RESULT 5
 ID AAF60372 standard; cDNA; 1813 BP.
 XX AC AAF60372;
 XX DT 27-APR-2001 (first entry)
 XX DE PRO246 coding sequence.
 XX KW Cytostatic; PRO protein; tumour; cancer; sa.
 XX OS Homo sapiens.
 XX PN W0200105836-A1.
 XX PD 25-JAN-2001.
 XX PF 20-DEC-1999; 99WO-US30999.
 XX PR 20-JUL-1999; 99US-0144758.
 XX PR 26-JUL-1999; 99US-0145698.
 XX PR 08-SEP-1999; 99WO-US20594.

301	agggtgtcttctccacgacctggagagtgcccttggatgtgggtcttccaacagaag	360	Db		1201	cccccatccctgggtgggttctctctctggcttgagcgcatgggtgtgctgctgca	1260
361	aaagagagatcagtggtgtctccatcaatctgggtgcaacaagcaaacctggagtat	420	Qy		1261	tgggtgctgccacagagtcgaagctgggtctctctggatgatgacccccaccactcattggcta	1320
361	aaagaggagatcaggtgtgtctccatcaatctgggtgcaacaagcaaacctggagtat	420	Db		1261	tgggtgctgccacagagtcgaagctgggtctctctggatgatgacccccaccactcattggcta	1320
421	ctctgggtactactcatcctccctccggaaacctgtccctgggtggaggtctccagaga	480	Qy		1321	aaggaatttgggttctctctctctataaaggatcacctctagcacagaggtcctgagtcag	1380
421	ctctgggtactactcatcctccctccggaaacctgtccctgggtggaggtctccagaga	480	Db		1321	aaggaatttgggttctctctctctataaagggtcacctctagcacagaggtcctgagtcag	1380
481	asgactctggccctacagctgtccgtgaatgcaagacaagagcaaatctagg	540	Qy		1381	ggaaagagtcacactcctgacaccttagtactctgccccacacctcttcttactgtgggaaa	1440
481	asgactctggccctacagctgtccgtgaatgcaagacaagagcaaatctagg	540	Db		1381	ggaaagagtcacactcctgacaccttagtactctgccccacacctcttcttactgtgggaaa	1440
541	gccacagcatcaaaaccttagaactcaatgtactgggttctccagctctccatctggc	600	Qy		1441	accatctcagtaagacctaaagttccaggagacagaagaggaagagagtgagatctgga	1500
541	gccacagcatcaaaaccttagaactcaatgtactgggttctccagctctccatctggc	600	Db		1441	accatctcagtaagacctaaagttccaggagacagaagaggaagagtgagatctgga	1500
601	gtctccagggtgtgcccatgtgggggcaaacgtgacctgagctgcagctctcaagga	660	Qy		1501	attgggaggagctccacccccctgactcctctatgaagccagctgctgaaataag	1560
601	gtctccagggtgtgcccatgtgggggcaaacgtgacctgagctgcagctctcaagga	660	Db		1501	attgggaggagctccacccccctgactcctctatgaagccagctgctgaaataag	1560
661	gtaagccgtgtccaataccagtgagatggcagcttccatcttccagactttcttg	720	Qy		1561	tactcacaagagtgaggggcagagacttcagtcactgagctctccaggcccccttga	1620
661	gtaagccgtgtccaataccagtgagatggcagcttccatcttccagactttcttg	720	Db		1561	tactcacaagagtgaggggcagagacttcagtcactgagctctccaggcccccttga	1620
721	caccagcattagatgcatacctgggtcttttaagcctcaccaaccttctgtcttcag	780	Qy		1621	tctgtacccccccttatcttaacacacaccttgggtccacactccagctccctgtattgat	1680
721	caccagcattagatgcatacctgggtcttttaagcctcaccaaccttctgtcttcag	780	Db		1621	tctgtacccccccttatcttaacacacaccttgggtccacactccagctccctgtattgat	1680
781	ctggagctcatgtctgaagccccaatgaagtgaggcactgcccaatgtaatgtgacgc	840	Qy		1681	ataactgtcaggctgggtgtgtttaggttttactggggcagaggatgggaatctcttat	1740
781	ctggagctcatgtctgaagccccaatgaagtgaggcactgcccaatgtaatgtgacgc	840	Db		1681	ataactgtcaggctgggtgtgtttaggttttactggggcagaggatgggaatctcttat	1740
841	tgggaagtgaacacagggcctgagctcagtggttctgtgagctgtgtgggtaccttg	900	Qy		1741	aaaaactaaactgaataatgtgttcttccatttgcgaacttcaataaagacacataatg	1800
841	tgggaagtgaacacagggcctgagctcagtggttctgtgagctgtgtgggtaccttg	900	Db		1741	aaaaactaaactgaataatgtgttcttccatttgcgaacttcaataaagacacataatg	1800
901	ttagactgggtgtcgtgctggctgggtcctcttgcaccacggcgggcaagggccttg	960	Qy				
901	ttagactgggtgtcgtgctggctgggtcctcttgcaccacggcgggcaagggccttg	960	Db				
961	aggagccagccaatgatatacaggagagatgccattgtctcccgagacctccttgccc	1020	Qy				
961	aggagccagccaatgatatacaggagagatgccattgtctcccgagacctccttgccc	1020	Db				
1021	agagctcagacacaatctccaagaatgggaccttctctgtcacctccgcacagagcc	1080	Qy				
1021	agagctcagacacaatctccaagaatgggaccttctctgtcacctccgcacagagcc	1080	Db				
1081	tccggccacccttgccctccagagctgggtgcatgacccccacggcctctctcca	1140	Qy				
1081	tccggccacccttgccctccagagctgggtgcatgacccccacggcctctctcca	1140	Db				
1141	gccagggcttgccttaccacagatgccacgacagatggggcccccacccaat	1200	Qy				
1141	gccagggcttgccttaccacagatgccacgacagatggggcccccacccaat	1200	Db				

[illegible]

22-FEB-2000; 2000NOV-US04414.

07-JUL-1999; 99US-0143048.
 26-JUL-1999; 99US-0145698.
 28-JUL-1999; 99US-0146222.
 08-SEP-1999; 99NOV-US20594.
 13-SEP-1999; 99NOV-US20944.
 15-SEP-1999; 99NOV-US21090.
 15-SEP-1999; 99NOV-US21547.
 05-OCT-1999; 99NOV-US23089.
 29-NOV-1999; 99NOV-US28214.
 30-NOV-1999; 99NOV-US28313.
 16-DEC-1999; 99NOV-US30095.
 20-DEC-1999; 99NOV-US30911.
 20-DEC-1999; 99NOV-US30999.
 05-JAN-2000; 99NOV-US00219.

(GETH) GENENTECH INC.

Ashtken AJ, Botstein D, Desnovers L, Eaton DL, Ferrara N;
 Pflavrovic E, Pong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
 Godowski PJ, Grimaldi C, Gurney AL, Hillan KJ, Klijavin LJ;
 Mather JP, Pan J, Paoi NF, Roy MA, Stewart TA, Tumas D;
 Williams PM, Wood WI;
 WPI; 2001-081051/09.
 P-PSDB; AAB80219.

Sixty one nucleic acids encoding PRO polypeptides which are useful in
 the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
 squamous cell carcinoma) and neurodegenerative diseases (e.g.
 Alzheimer's disease) -

Claim 2; Fig 16; 393pp; English.

The present sequence is one of sixty one nucleic acids encoding novel
 secreted and transmembrane PRO polypeptides. The PRO polypeptides are
 useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
 squamous cell carcinoma), gastrointestinal disorders (e.g.
 enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
 Parkinson's disease), wound repair, cardiovascular disorders (e.g.
 endometrial bleeding angiogenesis, ischaemia such as coronary
 ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
 rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
 diabetes and retinal disorders such as retinitis pigmentosa.
 The PRO nucleic acids have applications in molecular biology, including
 use as hybridization probes, and in chromosome and gene mapping.

Sequence 1813 BP; 368 A; 559 C; 484 G; 402 T; 0 other;

Query Match 100.0%; Score 1813; DB 22; Length 1813;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1813; Conservative 0; Mismatches 0; Indels 0; Gaps
 0;

1 ggagcgccctgggtgacggctcgagctccgcgcagctccggcgctcgagcgt 60
 |||||

Db 1801 ttgtgtgaaaaa 1813
 RESULT 8
 AAC97441
 ID AAC97441 standard; cDNA; 1813 BP.
 XX AAC97441;
 AC AAC97441;
 XX 28-FEB-2001 (first entry)
 XX Human angiogenesis-associated protein PRO246 cDNA, SEQ ID NO:95.
 DE Human; angiogenesis-associated protein; PRO; endothelial cell growth;
 KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;
 KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
 KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
 KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
 KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
 KW gene therapy; transgenic animal; ss.
 XX Homo sapiens.
 OS WO2000053753-A2.
 PW 14-SEP-2000.
 XX 05-JAN-2000; 2000WO-US00219.
 XX 08-MAR-1999; 99WO-US05028.
 PR 12-MAR-1999; 99US-0123957.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21947.
 PR 05-OCT-1999; 99WO-US23089.
 PR 30-NOV-1999; 99WO-US28113.
 PR 30-NOV-1999; 99WO-US28409.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 XX (GETH) GENENTECH INC.
 XX PI Abkhazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;
 PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;
 PI Piaci NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;
 DR WPI; 2001-090793/10.
 DR P-PSDB; AAB53082.

Qy	1021	agagctcagacacaatctcaagaatggagacctctctctgcacctccgcacagacc	1080
Db	1021	agagctcagacacaatctcaagaatggagacctctctctgcacctccgcacagacc	1080
Qy	1081	tcggccacccatggcctccaggctcgtgcatgacccccaagccagctctctcca	1140
Db	1081	tcggccacccatggcctccaggctcgtgcatgacccccaagccagctctctcca	1140
Qy	1141	gccaggccctgccctcaccaagatgccacgacagatggggcccccacccccaacat	1200
Db	1141	gccaggccctgccctcaccaagatgccacgacagatggggcccccacccccaacat	1200
Qy	1201	ccccatccccgggtgggggtttctctctggctgaagccgcatgggtgctgctctga	1260
Db	1201	ccccatccccgggtgggggtttctctctggctgaagccgcatgggtgctgctctga	1260
Qy	1261	tggtgctgcccagagtcgaagctggctctctggtatgatgaccccaacattggcta	1320
Db	1261	tggtgctgcccagagtcgaagctggctctctggtatgatgaccccaacattggcta	1320
Qy	1321	aaggatttggggctctctctctcataagggtcacctctagcacagagccctgagctg	1380
Db	1321	aaggatttggggctctctctctcataagggtcacctctagcacagagccctgagctg	1380
Qy	1381	ggaagagctcacactctgaccttagtactctgccccaacctctcttactgtggaaa	1440
Db	1381	ggaagagctcacactctgaccttagtactctgccccaacctctcttactgtggaaa	1440
Qy	1441	accatctcagtaagacctaaagtccaggagacagaaggagaaagtggatcttga	1500
Db	1441	accatctcagtaagacctaaagtccaggagacagaaggagaaagtggatcttga	1500
Qy	1501	attggggaggacctccaccacccctgacctctcttatgaagccagctgctgaaattg	1560
Db	1501	attggggaggacctccaccacccctgacctctcttatgaagccagctgctgaaattg	1560
Qy	1561	ctactcaccaagtgtaggggcagagactccagtcactgagctcccaggcccttga	1620
Db	1561	ctactcaccaagtgtaggggcagagactccagtcactgagctcccaggcccttga	1620
Qy	1621	tcctgaccaccaacctctataaccacaccttgctccactccagctccctgattgat	1680
Db	1621	tcctgaccaccaacctctataaccacaccttgctccactccagctccctgattgat	1680
Qy	1681	ataacctgcaggtgggttggttgggttactggggcagagatagggaattctctat	1740
Db	1681	ataacctgcaggtgggttggttgggttactggggcagagatagggaattctctat	1740
Qy	1741	taaaactaacatgaatatggtgttttctatttcgaaaatttaataaagatacataatg	1800
Db	1741	taaaactaacatgaatatggtgttttctatttcgaaaatttaataaagatacataatg	1800
Qy	1801	tttgtatgaaaaa 1813	
Db	1801	tttgtatgaaaaa 1813	

agents to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the polypeptides in samples (e.g. by enzyme linked immunosorbant assay (ELISA). Examples of diseases which may be treated include rheumatoid arthritis and diabetes.

Query Match	99.8%;	Score 1809;	DB 22;	Length 1821;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1809;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ggaagcccttgaggtggtcagcggtcggtcccgagcagctccgagcgtcgcgagcgt 60		
Db	12	ggaagcccttgaggtggtcagcggtcggtcccgagcagctccgagcgtcgcgagcgt 71		
Qy	61	cggacctcgaggtcgtcggtcccgcggtggcgccctgactcgctcccgccagggga 120		
Db	72	cggacctcgaggtcgtcggtcccgcggtggcgccctgactcgctcccgccagggga 131		
Qy	121	gggacatgatttcctcccggggcccctgggtgaccaactgctcggtttttgtctctcg 180		
Db	132	gggacatgatttcctcccggggcccctgggtgaccaactgctcggtttttgtctctcg 191		
Qy	181	gggtgagtgccctcgcgccctcccgggccagctgcaactgcactgcctcgcccaacc 240		
Db	192	gggtgagtgccctcgcgccctcccgggccagctgcaactgcactgcctcgcccaacc 251		
Qy	241	ggttgcaggcggtggaggagggaagtggtgtcttcagcgtggtacacttgcaagggg 300		
Db	252	ggttgcaggcggtggaggagggaagtggtgtcttcagcgtggtacacttgcaagggg 311		
Qy	301	aggtgtctcatccagccatgggaagtgcccttggtgtgtcttcacacagaaag 360		
Db	312	aggtgtctcatccagccatgggaagtgcccttggtgtgtcttcacacagaaag 371		
Qy	361	aaaaggaggtcaggtgtgtctcatcatcaatgggtgcacacgaacaaactggagat 420		
Db	372	aaaaggaggtcaggtgtgtgtctcatcatcaatgggtgcacacgaacaaactggagat 431		
Qy	421	ccttggtctactccatgccctcccggaacctgctcctcggtggagggtctccaggaga 480		
Db	432	ccttggtctactccatgccctcccggaacctgctcctcggtggagggtctccaggaga 491		
Qy	481	aagactcggccctacagctgctccggtgaatgtgaagacaacaaaggcaaatctaggg 540		
Db	492	aagactcggccctacagctgctccggtgaatgtgaagacaacaaaggcaaatctaggg 551		
Qy	541	gccacagatcaaaaccttagaactcaatgtactgtgtcttcacagctctcatctgcgc 600		
Db	552	gccacagatcaaaaccttagaactcaatgtactgtgtcttcacagctctcatctgcgc 611		
Qy	601	gtctccagggtgtgccccatgtgggggcaaacctgcacctgagctgccagttccaaggga 660		
Db	612	gtctccagggtgtgccccatgtgggggcaaacctgcacctgagctgccagttccaaggga 671		

XX The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension.
 XX
 SQ Sequence 1827 BP; 369 A; 559 C; 491 G; 408 T; 0 other;

Query Match 99.8%; Score 1806.6; DB 22; Length 1827;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1809; Conservative 0; Mismatches 4; Indels 0; Gaps

0;
 QY 1 ggagccgcccgtgggtgacggcgtcggtcccgccagcgtcgccgctcgccgagcgt 60
 DB 13 ggagccgcccgtgggtgacggcgtcggtcccgccagcgtcgccgctcgccgagcgt 72
 QY 61 cggcaactgcaggtcggtcggtcccgccggtcgcccgccagcgtcgccgagcgt 120
 DB 73 cggcaactgcaggtcggtcggtcccgccggtcgcccgccagcgtcgccgagcgt 132
 QY 121 ggcccatgattccctcccgccggtcggtgacaaactgctgcggtttttgtctctg 180
 DB 133 ggcccatgattccctcccgccggtcggtgacaaactgctgcggtttttgtctctg 192
 QY 181 ggctgagtgccctcgcccccctcgccgcccagctgcaactgcaactgcccgaacc 240
 DB 193 ggctgagtgccctcgcccccctcgccgcccagctgcaactgcaactgcccgaacc 252
 QY 241 ggctgagtgccctcgcccccctcgccgcccagctgcaactgcaactgcccgaacc 300
 DB 253 ggctgagtgccctcgcccccctcgccgcccagctgcaactgcaactgcccgaacc 312
 QY 301 aggtgtcttcatcccgccatggggagggagggtggtttccagcgtggtacacctgcacgg 360
 DB 313 aggtgtcttcatcccgccatggggagggagggtggtttccagcgtggtacacctgcacgg 372
 QY 361 aaaggaggatcaggtgtgtctctacatcaatgggtcacaaagcgaactggagat 420
 DB 373 aaaggaggatcaggtgtgtctctacatcaatgggtcacaaagcgaactggagat 432
 QY 421 ccttggttactactgacctcccggaacctgacctcggtggaggtgtccagagaga 480
 DB 433 ccttggttactactgacctcccggaacctgacctcggtggaggtgtccagagaga 492
 QY 481 agagcttggtccctcacgtgtcctggaatgacagagacaagaagcaactctagg 540
 DB 493 agagcttggtccctcacgtgtcctggaatgacagagacaagaagcaactctagg 552
 QY 541 gccacagatcaaaaccttagaactcaatgtactggttctccagctctccatctgccc 600
 DB 553 gccacagatcaaaaccttagaactcaatgtactggttctccagctctccatctgccc 612
 QY 601 gtctccagggtgtgccccatgtggtgggggcaaacgtgacctgagctgcccagctctccaagga 660

QY 1561 ctactaccagaagtgaggggcagagacttccagctactgagttcccgagccccccttga 1620
 DB 1572 ctactaccagaagtgaggggcagagacttccagctactgagttcccgagccccccttga 1631
 QY 1621 tctgtaccacccctatctaaacacacacttggttccactccagctccctgtattgat 1680
 DB 1632 tctgtaccacccctatctaaacacacacttggttccactccagctccctgtattgat 1691
 QY 1681 ataacctgcagctggcttggttaggttttactggcgagagatagggaactcttcat 1740
 DB 1692 ataacctgcagctggcttggttaggttttactggcgagagatagggaactcttcat 1751
 QY 1741 taaacaacatgaatgtgtgttttcttcatgttgcgaatttaataaagatacataatg 1800
 DB 1752 taaacaacatgaatgtgtgttttcttcatgttgcgaatttaataaagatacataatg 1811
 QY 1801 ttgtgatga 1809
 DB 1812 ttgtgatga 1820

RESULT 10

AAH02949

ID AAH02949 standard; DNA; 1827 BP.

XX AC AAH02949;

XX 15-JUN-2001 (first entry)

XX Human shear stress-response coding sequence SEQ ID NO: 143.

XX Human; shear stress-response protein; vascular disease;

XX arteriosclerosis; db.

XX Homo sapiens.

XX WO200125427-A1.

XX 12-APR-2001.

XX 02-OCT-2000; 2000WO-JP06840.

XX 01-OCT-1999; 99JP-0280976.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX (NOJ) NOJIMA H.

XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;

XX Kuga T, Sekine S, Nakamura Y, Sugano S;

XX WPI; 2001-266308/27.

XX P-PSDB; AAB90818.

XX DNA sequences, proteins encoded by them and antibodies against them

XX useful in diagnosis and treatment of vascular disease caused by

XX arteriosclerosis -

[illegible]

12-JUL-2001.

28-DEC-2000; 2000WO-JP09359.

06-JAN-2000; 2000JP-0000585.

06-JAN-2000; 2000JP-0000588.

11-JAN-2000; 2000JP-0002299.

03-FEB-2000; 2000JP-0026862.

03-MAR-2000; 2000JP-0058367.

(PROT-) PROTEGENE INC.

(SAGA) SAGAMI CHEM RES CENT.

Kato S, Kimura T;

WPI; 2001-418355/44.

P-PSDB; AAE06610.

Human proteins with hydrophobic domains and the nucleic acids encoding

Alzheimer's and inflammation

Claim 4; Page 486-489; 563pp; English.

The present sequence is human protein with hydrophobic domain encoding

cDNA clone HPI0801. The polynucleotide and polypeptide of the invention

may be used in the prevention, diagnosis and treatment of diseases

associated with inappropriate polypeptide expression. The

polynucleotides

may be used to produce the polypeptide, by inserting the nucleic acids

into a host cell and culturing the cell to express the protein. The

polynucleotides and its complementary sequences may also be used as DNA

probes in diagnostic assays and also used in gene therapy. The

polypeptides may also be used as antigens in the production of

antibodies

and in assays to identify modulators of polypeptide expression and

activity. The polypeptides and nucleic acids may be used as nutritional

supplements, to modulate cytokine and cell proliferation activity, to

modulate immune stimulation or suppression (e.g. for the treatment of

microbial infections and autoimmune disorders such as multiple

sclerosis,

rheumatoid arthritis and insulin-dependent diabetes), to modulate

hematopoiesis, to modulate tissue growth activity (e.g. for the

treatment of Parkinson's disease, Huntington's disease and Alzheimer's

disease), to modulate activin and inhibin activity (e.g. for controlling

fertility), to modulate chemotactic and chemokinetic activity, to

modulate haemostatic and thrombolytic activity, to modulate receptor

ligand activity, to modulate inflammation and to inhibit tumour growth.

Sequence 1816 BP; 362 A; 560 C; 488 G; 406 T; 0 other;

Query Match 99.5%; Score 1804.8; DB 22; Length 1816;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1806; Conservative 0; Mismatches 2; Indels 0; Gaps

0;

41

41

CC The invention relates to 40 human secreted proteins (AA994981-195020),
CC and cDNA sequences encoding them (AA923423-A93462). The secreted
CC proteins of the invention include those that are thought to be only
CC partially secreted, i.e., transmembrane proteins. The proteins of the
CC invention may exhibit one or more activities selected from the
following:

cytokine activity; cell proliferation; differentiation; immune modulation; haematopoiesis regulation; tissue growth activity; activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic and thrombolytic activity; anti-inflammatory activity; and tumour inhibition activity. The proteins may be administered to patients as vaccines, and the nucleotides may be used as part of a gene therapy regime. Diseases or conditions that may be treated using the proteins or nucleotides of the invention include autoimmune diseases; genetic disorders; haemophilia; cardiovascular diseases; cancer; bacterial, fungal and viral infections, especially HIV; multiple sclerosis; rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome; insulin dependent diabetes mellitus, and allergic reactions such as asthma and anaemia. They may also be used for treating wounds, burns, ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's disease, Parkinson's disease, Huntington's disease and amyotrophic lateral sclerosis (ALS). Proteins with activin/inhibin activity may additionally be useful as contraceptives. Nucleic acid sequences of the invention may be used in chromosome mapping, and as a source of diagnostic primers and probes. The present sequence represents cDNA encoding one of the 40 proteins of the invention.

Sequence 1954 BP; 498 A; 561 C; 490 G; 405 T; 0 other;

Query Match 99.4%; Score 1802; DB 21; Length 1954;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1813; Conservative 0; Mismatches 0; Indels 1; Gaps

1:

Qy 1 qqagcccccctgggtgtcagc-qqctcggctccgagcagctcggccagcc 59

11

Db 13 aaagccgccctgaattatcagcgggctcagctccggccacagctcggcccatcggccagcc 72

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

[illegible][illegible][illegible]

bp
73 cccgacacctgcaggctccgtcgcgtccccgccggctcgccccctgactccgtccccgccagg 137

[illegible]

oy 120 agggccatgatctccctccccgggcccctggcgaccaaacttgctgcgggttttgcctcctg 17

Db 133 agggccatgattccctcccgggccctggtagcaacttgctggggtttttgttcctg 192

Qy 180 gggctgagtgcctcgccccctcgccccctgcactgccaac 239

[illegible]

Db 193 gggctgagtgcctcgcccccctcgggccagctgcaactgcactgcccgcac 252

240 cggttgcaggcgtggaggagggaagtggctccagcgtggtacaccttgcacggg 299

db 253 cgggttcaggcggtagggagggaagtggtagctccaggcgtggtagacaccttgcacggg 312

.....

2y 300 gaggtatcttcattccacgccatgagagagatgccctttgtgatatgtgattcttcaaacagaaa 359

[illegible]

— 2 —

360	QY	gaaaaggaggatcaaggtgtgtctctacatcaatgggggtcacaacaaagcaaacctggagta	419
373	Db	gaaaaggaggatcaggtgtgtctctacatcaatgggggtcacaacaaagcaaacctggagta	432
420	QY	tcttggcttactctcatgccctccggaaacctgtccctggcggtggagggtctccaggag	479
433	Db	tcttggcttactctcatgccctccggaaacctgtccctggcggtggagggtctccaggag	492
480	QY	aaagactctggccctccacagctctccgtgaatgtgcaagacaaacaaagcgaatctagg	539
493	Db	aaagactctggccctccacagctctccgtgaatgtgcaagacaaacaaagcgaatctagg	552
540	QY	ggccacagcatcaaaacctagaactcaatgtactgtgtctccagctctccatctctgc	599
553	Db	ggccacagcatcaaaacctagaactcaatgtactgtgtctccagctctccatctctgc	612
600	QY	cgctctccaggggtgtgcccatgtggggggcaaacgtgacctgagctgccagctctccagg	659
613	Db	cgctctccaggggtgtgcccatgtggggggcaaacgtgacctgagctgccagctctccagg	672
660	QY	agtaagcccgctgtccataaccagtgggatcggcagcttccatctctccagactctctt	719
673	Db	agtaagcccgctgtccataaccagtgggatcggcagcttccatctctccagactctctt	732
720	QY	gcaccagcatagatgtcatccgtgggtctttaagcctcacacaacctcttgcgtctccatg	779
733	Db	gcaccagcatagatgtcatccgtgggtctttaagcctcacacaacctcttgcgtctccatg	792

[illegible]

42

[illegible]

CC secreted proteins.
 XX
 50 Sequence 1932 BP; 405 A; 585 C; 514 G; 424 T; 4 other;
 Query Match 98.4%; Score 1783.8; DB 21; Length 1932;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1804; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

QY 4 gccccttggtgtcagcgctcggtcccgccagcgtccggtcgccgagctcgg 63
 DB 9 gccccttggtgtcagcgctcggtcccgccagcgtccggtcgccgagctcgg 67

QY 64 caactcgaagtcgtcggtcccggtggtggtcccttgactccgtcccgccagggagg 123
 DB 68 caactcgaagtcgtcggtcccggtggtggtcccttgactccgtcccgccagggagg 127

QY 124 ccaatgattccctccggggccctggtgacaaactggtgcggtttttgttcctggggc 183
 DB 128 ccaatgattccctccggggccctggtgacaaactggtgcggtttttgttcctggggc 187

QY 184 tgagtgcctcgccccctccggtccggtccagctgcaactgcaactcccgccaaacgg 243
 DB 188 tgagtgcctcgccccctccggtccggtccagctgcaactgcaactcccgccaaacgg 247

QY 244 tgcaggcggtgagggagggaagtgtgctccagcgtggtacacttcgacggggagg 303
 DB 248 tgcaggcggtgagggagggaagtgtgctccagcgtggtacacttcgacggggagg 307

QY 304 tgtcttcacccagcgcattggaggtgcctttgtgattgtgtttctcaaacagaaagaa 363
 DB 308 tgtcttcacccagcgcattggaggtgcctttgtgattgtgtttctcaaacagaaagaa 367

QY 364 aggaggatcagtggtgtcctacatcaatgggggtcacaacagcaaacctggagttcct 423
 DB 368 aggaggatcagtggtgtcctacatcaatgggggtcacaacagcaaacctggagttcct 427

QY 424 tgggtctactccatgcctccggaaactgtccctgggtgaggggtccagggagaaag 483
 DB 428 tgggtctactccatgcctccggaaactgtccctgggtgaggggtccagggagaaag 487

QY 484 actctgccccctacagtcgtcctcgtaagtgcagagcaaaacaggaatctaggggcc 543
 DB 488 actctgccccctacagtcgtcctcgtaagtgcagagcaaaacaggaatctaggggcc 547

QY 544 acagatcaaaactctagaactcaatgactggttctccagctctccatctccgtcc 603
 DB 548 acagatcaaaactctagaactcaatgactggttctccagctctccatctccgtcc 607

QY 604 tccagggtgtgccccatgtgggggcaaacgtgacctgagctgacctcccaaggagta 663
 DB 608 tccagggtgtgccccatgtgggggcaaacgtgacctgagctgacctcccaaggagta 667

QY 664 agcccgtgtccaaatccagctggagtcggcagcttccatctccagacttttttgac 723
 DB 668 agcccgtgtccaaatccagctggagtcggcagcttccatctccagacttttttgac 727

DB 728 cagcattagatgcattccggtgttttaagctcaccacacttctgtcttccatggctg 787

QY 784 gagtctatgtctcgaagccccaatgagtggtgggcaactgccaatgtaatgtgacgtgg 843
 DB 788 gagtctatgtctcgaagccccaatgagtggtgggcaactgccaatgtaatgtgacgtgg 846

QY 844 aagtgtgacacagggccttgagctgcagtcagtggtgtgtgagctgtgtgggtacccgtg 903
 DB 847 aagtgtgacacagggccttgagctgcagtcagtggtgtgtgagctgtgtgggtacccgtg 906

QY 904 gactgggggtgctgggtgggtggtcctctgtaccacccggggcaaggccctggagg 963
 DB 907 gactgggggtgctgggtgggtggtcctctgtaccacccggggcaaggccctggagg 966

QY 964 agccagccaatgatcaaggaggatgccattgctccccgggacctgccccgtgccaaga 1023
 DB 967 agccagccaatgatcaaggaggatgccattgctccccgggacctgccccgtgccaaga 1026

QY 1024 gtcagacacaatctccaagaatgggacctttctctgtcactccgcagcagcctcc 1083
 DB 1027 gtcagacacaatctccaagaatgggacctttctctgtcactccgcagcagcctcc 1086

QY 1084 ggccacccatggccctcccaagcctgggtgcattgacccccacgcgcagctctccagcc 1143
 DB 1087 ggccacccatggccctcccaagcctgggtgcattgacccccacgcgcagctctccagcc 1146

QY 1144 aggcctctgctctcaccagaactgccacgacagatggggccaccctcaaccaatctccc 1203
 DB 1147 aggcctctgctctcaccagaactgccacgacagatggggccaccctcaaccaatctccc 1206

QY 1204 ccactccctggtgggggtttctctctggtgagcgcagtggtgctgctgctgtgatgg 1263
 DB 1207 ccactccctggtgggggtttctctctggtgagcgcagtggtgctgctgctgtgatgg 1266

QY 1264 tgcctgccagagtcgaagtcggctctctggtatgatgacccccaccactcattggctaaag 1323
 DB 1267 tgcctgccagagtcgaagtcggctctctggtatgatgacccccaccactcattggctaaag 1326

QY 1324 gatttgggtgtctccttcctataaagggtcacctctagcacagagcctgagtcagggga 1383
 DB 1327 gatttgggtgtctccttcctataaagggtcacctctagcacagagcctgagtcagggga 1386

QY 1384 aagagtcacactcctgacctttagtactctgcccccaactctcttactgtgggaaaaacc 1443
 DB 1387 aagagtcacactcctgacctttagtactctgcccccaactctcttactgtgggaaaaacc 1446

QY 1444 atctcagttaagccttaagtcacaggacagagaaggaagagagtggaatcggaaatc 1503
 DB 1447 atctcagttaagccttaagtcacaggacagagaaggaagagagtggaatcggaaatc 1506

QY 1504 gggaggagcctccaccacccctgactcctctctatgaagccagctgtgtaaatagctga 1563
 DB 1507 gggaggagcctccaccacccctgactcctctctatgaagccagctgtgtaaatagctga 1566

QY 1564 ctccacaagagtgaggggcagagactccagtcactgagtcctccagggcccccttgatct 1623

CC and polypeptides encoded by the genes expression vectors and must be used for producing the polypeptides are disclosed and methods for screening or purifying ligands which specifically bind to the polypeptides are also provided. The polynucleotides are useful for treating diseases associated with the altered expression of a gene that is coexpressed with one or more known atherosclerosis-associated genes in a subject. They are useful in diagnosis, prognosis, treatment, prevention, selection and evaluation of therapies for atherosclerosis including stroke, myocardial infarction, transient cerebral ischemia, mesenteric ischemia, coronary artery disease, angina pectoris, peripheral vascular disease, renal artery stenosis, and hypertension. Sequences AAC5065-85098 represent CC atherosclerosis-associated genes of the invention.

XX
SQ Sequence 1831 BP; 370 A; 561 C; 494 G; 406 T; 0 other;

Query Match 97.1%; Score 1760.4; DB 22; Length 1811;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 6; Indels 13; Gaps 3;

Qy 1 ggagccgcccgggtgacggcgtcggctccgcgcacgtccggcgtcgccgcgcgcct 60
| | | | |
Db 3 ggagccgcccgggtgacggcgtcggctccgcgcacgtccggcgtcgccgcgcct 62
| | | | |
Qy 61 cggacactgcaggtccgt-gcgcccgccggt-gggcccccgtcgtccgcgcgcag 118
| | | | |
Db 63 ggagccgcccgggtgacggcgtcggctccgcgcacgtccggcgtcgccgcgcag 122
| | | | |
Qy 119 gaggccatgatctccctccgggcccctggtagcaactgctgcggtttgttctct 178
| | | | |
Db 123 gaggccatgatctccctccgggcccctggtagcaactgctgcggtttgttctct 182
| | | | |
Qy 179 ggggtgagtgccctcgcccccctcgccgcccagctgcaactgcaacttccgcgcaa 238
| | | | |
Db 183 ggggtgagtgccctcgcccccctcgccgcccagctgcaactgcaacttccgcgcaa 242
| | | | |
Qy 239 ccggtgagtcggcgtggaggagggaagtgggtcttcacagcgtggtacacttgcacgg 298
| | | | |
Db 243 ccggtgagtcggcgtggaggagggaagtgggtcttcacagcgtggtacacttgcacgg 302
| | | | |
Qy 299 ggaagtgtcttcacccagcgcgtggaggtgccccttggatggtgttctcaacagaa 358
| | | | |
Db 303 ggaagtgtcttcacccagcgcgtggaggtgccccttggatggtgttctcaacagaa 362
| | | | |
Qy 359 ggaagaggagatcaggtgtgtctacatcaatgggtcacaacagcaacccctggagt 418
| | | | |
Db 363 ggaagaggagatcaggtgtgtctacatcaatgggtcacaacagcaacccctggagt 422
| | | | |
Qy 419 atcttggcttactcactccctccggaacctgtccctcggtcgagggtgtccagga 478
| | | | |
Db 423 atcttggcttactcactccctccggaacctgtccctcggtcgagggtgtccagga 482
| | | | |
Qy 479 gaaagacttggccctcagctcgtcgatgtaagtgcagacaaacagcaaatctag 538
| | | | |
Db 483 gaaagacttggccctcagctcgtcgatgtaagtgcagacaaacagcaaatctag 542
| | | | |

.Db 1567 ctaccagagtgaggggcagagacttccagctcagtgctccaggcccccctgtatt 1626
| | | | |
Qy 1624 gtaccccccctctatcaacacaccccttggctccactccagctccctgtattgata 1683
| | | | |
Db 1627 gtaccccccctctatcaacacaccccttggctccactccagctccctgtattgata 1686
| | | | |
Qy 1684 acctgcaggctggctggttaggtttactggggcagaggtagggaatctcttataa 1743
| | | | |
Db 1687 acctgcaggctggctggttaggtttactggggcagaggtagggaatctcttataa 1746
| | | | |
Qy 1744 aactaacatgaatatggtgtgttttcatttgcaaatataaagacacataatgttt 1803
| | | | |
Db 1747 aactaacatgaatatggtgtgttttcatttgcaaatataaagacacataatgttt 1806
| | | | |
Qy 1804 gtagtaaaaa 1813
| | | | |
Db 1807 gtagtaaaaa 1816
| | | | |

RESULT 14
AAC5076
ID AAC5076 standard; DNA; 1831 BP.
XX AAC5076;
XX
XX
DT 08-MAY-2001 (first entry)
XX
DE Atherosclerosis-associated gene seq ID No. 12.
XX
KW Atherosclerosis-associated gene; stroke; myocardial infarction; human;
KW ischemia; coronary artery disease; angina pectoris; hypertension;
KW peripheral vascular disease; renal artery stenosis; atherosclerotic;
KW cerebroprotective; cardiac; gene therapy; hypotensive; vasotropic;
KW antianginal; ds.
XX
XX Homo sapiens.
XX
FN WO200104264-A2.
XX
XX 18-JAN-2001.
XX
XX 28-JUN-2000; 2000WO-US17887.
XX
XX 07-JUL-1999; 99US-0349015.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Jones KA, Volkmut W, Walker MG;
XX
XX WPI; 2001-138330/14.
XX
XX Composition comprising atherosclerosis-associated polynucleotide useful
XX in diagnosis, prognosis, treatment, and prevention of atherosclerosis
XX and stroke, myocardial infarction, or hypertension -
XX
XX Claim 1; Page 43; 58pp; English.
XX
XX The invention provides novel atherosclerosis-associated polynucleotides
CC

Db	543	gggcccagcatcaaaacctagaactcaatgactgggttctccagctctccatctgt	602
Qy	599	ccgtctccagggtgtgcccatgtgggggcaaacgtgaccctgagctgcccagctccaag	658
Db	603	ccgtctccagggtgtgcccatgtgggggcaaacgtgaccctgagctgcccagctccaag	662
Qy	659	gagtaagccgcgtgtccaataccgtgggatggcagcttccatctccagactttctt	718
Db	663	gagtaagccgcgtgtccaataccgtgggatggcagcttccatctccagactttctt	722
Qy	719	tgccacgacattagatgtcatcggtgtcttaagcctcaccacatttctgtctccat	778
Db	723	tgccacgacattagatgtcatcggtgtcttaagcctcaccacatttctgtctccat	782
Qy	779	ggctgagctctatgtctgaagggcccaaatgaggtgggcaatgcccaatgaatgtgac	838
Db	783	ggctgagctctatgtctgaagggcccaaatgaggtgggcaatgcccaatgaatgtgac	842
Qy	839	gctgggaagtgaacaca-----gggctggagctgcagctggtgtgagctgtt	887
Db	843	gctgggaagtgaacaca-----gggctggagctgcagctggtgtgagctgtt	902
Qy	888	gtgggtaccctgggtgagctgggtgtgctgggtgggtgtgctctgtacacacgcgg	947
Db	903	gtgggtaccctgggtgagctgggtgtgctgggtgggtgtgctctgtacacacgcgg	962
Qy	948	ggcaaggccctggaggagcagcaatgatcaaggagagtgccattgctcccgagcc	1007
Db	963	ggcaaggccctggaggagcagcaatgatcaaggagagtgccattgctcccgagcc	1022
Qy	1008	ctgcccctggcccaagagctcagacacaaatctcaagaatgggaccccttctctgtcacc	1067
Db	1023	ctgcccctggcccaagagctcagacacaaatctcaagaatgggaccccttctctgtcacc	1082
Qy	1068	tccgcagcagccctccggccaccatggccctccagggctgggtgcattgacccccacg	1127
Db	1083	tccgcagcagccctccggccaccatggccctccagggctgggtgcattgacccccacg	1142
Qy	1128	cccagctctccagcaggccctgcctcaccagactgcccaacagatggggccac	1187
Db	1143	cccagctctccagcaggccctgcctcaccagactgcccaacagatggggccac	1202
Qy	1188	ctcaacaaatcccccatccctgggtgggttttctctgggtgagccgcagtggt	1247
Db	1203	ctcaacaaatcccccatccctgggtgggttttctctgggtgagccgcagtggt	1262
Qy	1248	gctgtgctgtatggtgctccagagtcgaagctggctctctggatgatgacccac	1307
Db	1263	gctgtgctgtatggtgctccagagtcgaagctggctctctggatgatgacccac	1322
Qy	1308	cactcattggctaaagatttgggtgtctctctcctataagggtcactctagcacaga	1367
Db	1323	cactcattggctaaagatttgggtgtctctctcctataagggtcactctagcacaga	1382
Qy	1368	ggcctgagctcagggaaagagtcacactcctgacccttagtactgtccccactctct	1427
Qy	1428	ttactgtgggaaaccatctcagtaagacactaagtgtccaggagacagaaggaagagg	1487
Db	1443	ttactgtgggaaaccatctcagtaagacactaagtgtccaggagacagaaggaagagg	1502
Qy	1488	aagtggatctggaattgggagagcctccaccacccctgactctctcttatgaagccag	1547
Db	1503	aagtggatctggaattgggagagcctccaccacccctgactctctcttatgaagccag	1562
Qy	1548	ctgctgaaatagctactcaccagagtgaggggcagagacttccagtcactgagctcc	1607
Db	1563	ctgctgaaatagctactcaccagagtgaggggcagagacttccagtcactgagctcc	1622
Qy	1608	caggcccccttgactgtgctaccccccactatctaacacacccctggctccactccagc	1667
Db	1623	caggcccccttgactgtgctaccccccactatctaacacacccctggctccactccagc	1682
Qy	1668	tccctgtattgatataacctgtcagggtggctgttaggttttactgggggcagagata	1727
Db	1683	tccctgtattgatataacctgtcagggtggctgttaggttttactgggggcagagata	1742
Qy	1728	gggaatctcttataaactaaactgaatgaatgtgtgttttcttgcataattaaata	1787
Db	1743	gggaatctcttataaactaaactgaatgaatgtgtgttttcttgcataattaaata	1802
Qy	1788	aagatacataatgtttgtatgaaaaa	1813
Db	1803	aagatacataatgtttgtatgagata	1828

RESULT 15
 AAF44978
 ID AAF44978 standard; cDNA; 1869 BP.
 XX AAF44978;
 XX 28-MAR-2001 (first entry)
 XX Human INTERCEPT 258 coding sequence SEQ ID NO. 26.
 Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
 INTERCEPT 258; coronary disorder; olfactory disorder;
 neurological disorder; pulmonary disorder; immunological disorder;
 developmental disorder; kidney disorder; ss.
 Homo sapiens.
 WO200078808-A1.
 28-DEC-2000.
 19-JUN-2000; 2000WO-US16883.
 18-JUN-1999; 99US-0336536.
 (MILL-) MILLENNIUM PHARM INC.
 PA
 XX

QY 1379 tgggaagagtcacactcctgacccttagtactctgccccacactctcttactgtggga 1438
 Db |||||
 QY 1408 tgggaagagtcacactcctgacccttagtactctgccccacactcttactgtggga 1467
 Db |||||
 QY 1439 aaacatctcagtaagacctaagtgtccaggagacagaaggaaggaagtgatctg 1498
 Db |||||
 QY 1468 aaacatctcagtaagacctaagtgtccaggagacagaaggaaggaagtgatctg 1527
 Db |||||
 QY 1499 gaattggaggagcctcacccacccctgactcctcttataagccagctgctgaatt 1558
 Db |||||
 QY 1528 gaattggaggagcctcacccacccctgactcctcttataagccagctgctgaatt 1587
 Db |||||
 QY 1559 agctactcaccaagagtgaaggagagactccagtcactgagctccaggccctt 1618
 Db |||||
 QY 1588 agctactcaccaagagtgaaggagagactccagtcactgagctccaggccctt 1647
 Db |||||
 QY 1619 gatctgacccacccctatctaaacacaccccttggtccactccagctccctgattg 1678
 Db |||||
 QY 1648 gatctgacccacccctatctaaacacaccccttggtccactccagctccctgattg 1707
 Db |||||
 QY 1679 atataacctgcaggctggctgggtagggttttactggggcagaggatagggaactctt 1738
 Db |||||
 QY 1708 atataacctgcaggctggctgggtagggttttactggggcagaggatagggaactctt 1767
 Db |||||
 QY 1739 attaaactaacatgaaatagtgtgtgttttcatttgcaatttaataaagatacataa 1798
 Db |||||
 QY 1768 attaaactaacatgaaatagtgtgttttcatttgcaatttaataaagatacataa 1827
 Db |||||
 QY 1799 tgtttgtatgaaaaa 1813
 Db |||||
 QY 1828 tgtttgtatgagata 1842
 Db |||||

Search completed: August 19, 2002, 16:16:06
 Job time: 4314 sec

OM nucleic - nucleic search, using sw model

Run on: August 19, 2002, 15:02:47 ; Search time 54.72 Seconds
(without alignments)

8138.405 Million cell updates/sec

Title: US-09-902-759-38

Perfect score: 1813

Sequence: 1 gggagccctgggtgag.....cataatgtttgtatgaaaaa 1813

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA.*
- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
 - 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
 - 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
 - 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
 - 5: /cgn2_6/ptodata/2/ina/6C_COMB.seq.*
 - 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1376	75.9	1387	2	US-08-979-424-2
2	51.4	2.8	7218	1	US-08-232-463-14
3	45.4	2.5	1584	4	US-08-928-383B-1
4	45.4	2.5	2434	4	US-09-272-496-1
5	44.2	2.4	4403765	4	US-09-103-840A-2
6	43.4	2.4	4403765	4	US-09-103-840A-2
7	42.4	2.3	1095	4	US-08-928-383B-3
8	42.2	2.3	2830	1	US-07-882-292-1
9	42.2	2.3	2830	2	US-08-331-644-1
10	42.2	2.3	2830	5	PT-US93-04102-1
11	41.8	2.3	477	4	US-09-135-994-1
12	41.6	2.3	1515	4	US-08-928-383B-25

ALIGNMENTS

RESULT 1
US-08-979-424-2
; Sequence 2, Application US/08979424
; Patent No. 5942606
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: VIRAL RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS


```

Q# .1261 tgggtcctgccagagtcgaagtgcctctctggtatgatcccccacactcatggcta 1320
Db 1260 TGGTGTCTCTGCCAGAGTCGAAGCTGGCTCTCTGGGTATGATGACCCACCACATCATGGCTA 1319
Q# 1321 agggatttgggtctctctctctataagggtcacctctagcacagagcgctgagtcagt 1380
Db 1320 AGGATTGGGGTGCTCTCTCTCTATAAGGTCATCTCTAGCACAGAGGCCTGATGTCATG 1379
Q# 1381 ggaaagag 1388
Db 1380 GGAAGAAG 1387

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: PALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6

```

ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

OM nucleic - nucleic search, using sw model

Run on: August 19, 2002, 15:00:57 ; Search time 1589.67 Seconds
(without alignments)
15393.119 Million cell updates/sec

Title: US-09-302-759-38
Perfect score: 1013
Sequence: 1 ggcgcgcctgggtcgag.....cataaagtgtgatgaaaaa 1813

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- EST:*
 - 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_gss:*
 - 13: em_gss_hum:*
 - 14: em_gss_inv:*
 - 15: em_gss_pln:*
 - 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description

C 1	1026.2	56.6	1069	9	AL573851
2	1005.6	55.5	1017	9	AL547358

AL550211	AL550211	5	947.6	52.3	1013	9	AL550211
AL552901	AL552901	6	942.6	52.0	1003	9	AL552901
AL571713	AL571713	C 7	942	52.0	955	9	AL571713
AL573957	AL573957	C 8	937.2	51.7	1026	9	AL573957
AL548482	AL548482	C 9	911.6	50.3	969	9	AL548482
AL549789	AL549789	10	911	50.2	1109	9	AL549789
AL572321	AL572321	C 11	910.2	50.2	998	9	AL572321
AL549049	AL549049	C 12	906.6	50.0	973	9	AL549049
AL576327	AL576327	C 13	896	49.4	978	9	AL576327
AL546335	AL546335	14	895.8	49.4	938	9	AL546335
AL554790	AL554790	15	895.8	49.4	999	9	AL554790
AL574910	AL574910	C 16	893.6	49.3	938	9	AL574910
AL577433	AL577433	C 17	890.6	49.1	999	9	AL577433
AL553280	AL553280	18	886.2	48.9	920	9	AL553280
AL551281	AL551281	19	883.6	48.7	946	9	AL551281
AL537139	AL537139	20	863.6	47.6	894	9	AL537139
AK009223	AK009223	21	843.6	46.5	1816	11	AK009223
MUSCU							
AL549957	AL549957	22	842.8	46.5	907	9	AL549957
BC019460	BC019460	23	842.4	46.5	1859	11	BC019460
MUSCU							
AL550946	AL550946	24	830.6	45.8	900	9	AL550946
BI771838	BI771838	25	819.4	45.2	853	10	BI771838
603055231							
AL555047	AL555047	26	809.8	44.7	853	9	AL555047
BI769077	BI769077	27	806.8	44.5	834	10	BI769077
603057054							
AL549197	AL549197	28	798	44.0	878	9	AL549197
AL550912	AL550912	C 29	787.6	43.4	886	9	AL550912
AL551438	AL551438	30	784.8	43.3	790	9	AL551438
BI767362	BI767362	31	777	42.9	885	10	BI767362
603057503							
BI821049	BI821049	32	766	42.3	789	10	BI821049
603035511							
AL577654	AL577654	C 33	764.4	42.2	795	9	AL577654
BI819730	BI819730	34	758.2	41.8	872	10	BI819730
603041421							
BI771884	BI771884	35	737.8	40.7	754	10	BI771884
BI916766	BI916766	36	726.2	40.1	774	10	BI916766
603178773							
AL547289	AL547289	37	723.6	39.9	735	9	AL547289
BI759939	BI759939	38	723.6	39.9	833	10	BI759939
603044368							
BI818810	BI818810	39	718	39.6	905	10	BI818810
603037664							
AL566674	AL566674	C 40	710	39.2	739	9	AL566674
BI772732	BI772732	41	707.4	39.0	753	10	BI772732
603053265							
BI759963	BI759963	42	703.8	38.8	760	10	BI759963
602733576							
BI909870	BI909870	43	699.8	38.6	782	10	BI909870
603070501							
BI758052	BI758052	44	690.4	38.1	748	10	BI758052
602712209							
BI517736	BI517736	45	666.2	36.7	861	10	BI517736

ALIGNMENTS

RESULT 1
 AL573851/c
LOCUS
 16-FEB-2001
DEFINITION AL573851 LTI_NFL006_P12 Homo sapiens cDNA clone CS0DI054YI01.3
 prime, mRNA sequence.
ACCESSION
 AL573851
VERSION
 AL573851.1 GI:12933489
KEYWORDS
 EST.
SOURCE
 human.
ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1. (bases 1 to 1069)
AUTHORS
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
 Full-length cDNA libraries and normalization
JOURNAL
 Unpublished (2001)
COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
 source
 1. 1069
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DI054YI01"
 /tissue_type="placenta"
 /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"
 BASE COUNT 248 a 269 c 327 g 224 t 1 others
 ORIGIN
 Query Match 56.6%; Score 1026.2; DB 9; Length 1069;
 Best Local Similarity 99.5%; Pred. No. 2.7e-232;
 Matches 1050; Conservative 0; Mismatches 3; Indels 2; Gaps
 2;
 Qy 714 ttctttgaccagcattagatgcatccgtggtgttttaagcctcaccacatttcgtct 773
 |||||
 Db 1061 tttcttgaccagcattagatgcatccgtggtgttttaagcctcaccacatttcgtct 1002
 |||||
 Qy 774 tccatggctggagctctatgtctgcaaggcccaaatgaggtgggcaatgtaata 833

QY	1633	ccctataataaccaccccttggtcccaactccagctccctggtatgataaccctgcag	1692
QY	11	tgggtgtcagcggtctgggtcccgccagcagctccggccgctcgccagcagctcggaacctgcg	70
Db	148	ccctattataacacaccccaatggctcccaacscgscscgkgtattgtatataacctgtcag	89
QY	1693	gctggctgttagcttttctcgggagagtagggaatcctctataaactaacat	1752
Db	88	gctggctgttagcttttctcgggagagtagggaatcctctataaactaacat	29
QY	1753	gaatatgtgtgttttctcatttgcacat	1780
Db	28	gaatatgtgtgttttctcatttgcacat	1
RESULT	5		
AL550211			
LOCUS	AL550211	1013 bp	linear
DEFINITION	AL550211 LTI_NFL006_P2 Homo sapiens cDNA clone CS01039YD07 5		EST
ACCSSION	AL550211		
VERSION	AL550211.1	GI:12886963	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1013)		
TITLE	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
FEATURES	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
source	1..1013		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="CS01039YD07"			
/clone_lib="LTI_NFL006_P2"			
/tissue_type="placenta"			
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371			
Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"			
BASE COUNT	183 a 309 c	304 g	216 t 1 others
ORIGIN			
Query Match	52.3%	Score 947.6;	DB 9; Length 1013;
Best Local Similarity	99.3%	Pred. No. 1.1e-213;	
Matches 1003;	Conservative 1;	Mismatches 0;	Indels 6;
Gaps			

Db 840 ACAGGCGCTGAGAGCTGCAGAGTGTCTGAGAGTGTCTGAGTGA-CTGTGAGAGTGGG 898
 Qy 912 ttgctggtggcgtggtctctttgtatcaccccgagggaagccctggagagcagcc 971
 Db 899 TTGCTGAGTGGGCTGGTCTCTTGTGA-CACCGCGGGGCAAGGCC--TGGAGAGCAACC 955
 Qy 972 aatgatatcaaggga 985
 Db 956 AATGATATHAGGA 969
 RESULT 10
 LOCUS AL549789 1109 bp mRNA linear EST
 DEFINITION AL549789 LTI_NFL006.PL2 Homo sapiens cDNA clone CSODI054YI01 5
 prime, mRNA sequence.
 ACCESSION AL549789
 VERSION AL549789.1 GI:12886119
 KEYWORDS EST
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (Bases 1 to 1109)
 AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
 TITLES Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source 1..1109
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODI054YI01"
 /clone_lib="LTI_NFL006.PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-oligo (dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/
 BASE COUNT 204 a 337 c 238 t 3 others
 ORIGIN
 Query Match 50.2%; Score 911; DB 9; Length 1109;
 Best Local Similarity 98.9%; Pred. No. 5.4e-205;
 Matches 937; Conservative 1; Mismatches 6; Indels 3; Gaps

Qy 10 ctgggtgtcagcggctcgggtcccgccgacgctccggcgtcgagccttcgggacctg 69
 Db 1 CTGGGTTGTCAGCGGCTCGGCTCCCGCGACGCTCCCGCGCTCGCGCA-CTCGGCACTG 59
 Qy 70 caggtccgtgctcccgcggtggcgccctgactccgtccggccagggagggccatga 129
 Db 60 CAGGTCGGTTCGCTCCCGCGGCTGGCGCCCTGACTCCGTCCCGGCGCAGGAGGCGCATGA 119
 Qy 130 ttctcctcccgggggccctcgtgaccacttctcggtttttctcgtgggctgagtg 189
 Db 120 TTTCCCTCCCGGGGCCCTCGTGACAACTGTCTGGGTGTTTTGTTCTCGGGGCTGAGTG 179
 Qy 190 ccttcgccccctccggtggcccgctgcaactgcaactgcccgcgaacccggtgcagg 249
 Db 180 CCCTCGGCCCCCTTCGGGGGCCAGCTGCACTGCACTTGCTGCCCAACCGGTTCGAGG 239
 Qy 250 cgtggaggaggaggaagtggctccagcgtgggtacacctgacccggggaggggtctct 309
 Db 240 CGGTGGAGGAGGGGAGTGGTCTCTCCGGGTGGTACACCTGTCACGGGGAGGTCTCT 299
 Qy 310 catccagcctcggaggtgcccctttgtgattggtgttcttcaacagaaagaaaggagg 369
 Db 300 CATCCAGCCATGGAGGTGCCCTTTGTGATGTGGTTCTTCMAACAGAAAGAAAGAGG 359
 Qy 370 atcaggtgttgcctcatcaatgggggtgcacacagcaaacactcggagtcaccttggctct 429
 Db 360 ATCAGGTGTTGCTTACATCATATGGGGTCAACAAGCAAGCAAACTTGGAGTATCTCTTGGTCT 419
 Qy 430 actcactgcccctccggaaactgctccctggggtgggggtggtccagggagaaagactcg 489
 Db 420 ACTCATGTCCTCCCGGAACTCTTCCTGGCGTGGAGGGGTCTCCAGGAGAAAGACTCTG 479
 Qy 490 gcccctacagctgctccgtgaatgtgcaagacaacaaggcaaatctagggggcccacagca 549
 Db 480 GCCCTACAGCTCTCCGTGAATGTGCAGACAAACAGGCAAAATCTAGGGGCCACAGCA 539
 Qy 550 tcaaaccttagaactcaatgactcgggttcctccagctcctccatcctcgcgtctccagg 609
 Db 540 TCMAAACCTTAGAAGTCAATGATGTTCTCTCCAGCTCTCCATCTCTCGCTCTCCAGG 599
 Qy 610 gtgtcccccatgtgggggaaacgtgacctcagctgagctgcccagtcctcaaggagtaagcccg 669
 Db 600 GTGTCCCCCATGTGGGGCAACGTCGACCTTGAGCTGCGAGTCTCCAGGAGTAAGCCCG 659
 Qy 670 ctgtccaataccagtgaggatcggcagcttcacatccttcagagattttttgaccacgat 729
 Db 660 CTGTCCAAATACCAAGTGGGATCGGAGCTTCCATCTCTTCAGCACTTCTTTTGACCCAGCAT 719
 Qy 730 tagatgctccgtgggtctttaagcctcaccaactcttcgtctcctcagtcggtgagctct 789
 Db 720 TAGATGTCATCCGTGGGTCTTTAAGGCTTCACCAACCTTTGCTCTTCATCGGCTGAGTCT 779
 Qy 790 atgtctgcaaggcccaacatgagtgagggtgggactgcccaatgtaatgagcgtggaagtga 849
 Db 780 ATGTCTGCMAGGCCCAACATGATGATGGGACTGCCCCCAATGATGATGAGCCCTGAGACTGA 839

63

QY	DB	100	GTCTCCCTGTTATGATATTAACCTGTTCAGGCTGTGGTGTAGGTTTATCTCCGGGCAGAGGA	41
QY	Db	1726	tagggaatctcttataaactaaacatgaataatgtgtgttt	1768
QY	Db	40	TAGGAACAC---TTATTAACATACAT-AAATATGTGTGTT	2
RESULT 12				
QY	AL549049	AL549049	973 bp	linear
QY	LOCUS	AL549049	973 bp	mrna
QY	16-FEB-2001	AL549049	973 bp	linear
QY	DEFINITION	AL549049 LTI_NFL006_P12 Homo sapiens cDNA clone CS0D1042YAL7 5		
QY	ACCESSION	AL549049	Prime, mRNA sequence.	
QY	VERSION	AL549049.1	GI:12884656	
QY	KEYWORDS	EST.		
QY	SOURCE	human.		
QY	ORGANISM	Homo sapiens		
QY	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
QY	AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayres, D.		
QY	JOURNAL	Full-length cDNA libraries and normalization		
QY	COMMENT	Unpublished (2001)		
QY	CONTACT	Contact: Genoscope		
QY	Genoscope - Centre National de Sequencage			
QY	BP 191 91006 EVRY cedex - France			
QY	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.			
QY	FEATURES	Location/Qualifiers		
QY	source	1..973		
QY	/organism="Homo sapiens"			
QY	/db_xref="taxon:9606"			
QY	/clone="CS0D1042YAL7"			
QY	/clone_lib="LTI_NFL006_P12"			
QY	/tissue_type="placenta"			
QY	/note="Vector: PCWVSPORT 6; Site_1: NotI; 1st strand cDNA			
QY	was primed with a NotI-oligo(dT) primer. Five prime end			
QY	enriched, double-stranded cDNA was digested with Not I and			
QY	cloned into the Not I and Eco RV sites of the PCWVSPORT 6			
QY	vector. Library was normalized. Library was constructed by			
QY	Life Technologies. Contact : Feng Liang Life Technologies,			
QY	a division of Invitrogen 9800 Medical Center Drive			
QY	Rockville, Maryland 20850, USA Fax : (1) 301 610 8371			
QY	Email : fliang@lifetech.com URL :			
QY	http://fulllength.invitrogen.com"			
QY	BASE COUNT	177 a 289 g 209 t		
QY	OTHER	2 others		
QY	ORIGIN			
QY	Query Match	50.0%	Score 906.6; DB 9; Length 973;	
QY	Best Local Similarity	98.6%	Pred. No. 5.6e-204;	
QY	Matches 965; Conservative	2; Mismatches	6; Indels	6; Gaps
QY	5;			
QY	13	gggtgcagcggtcgagctccgcgcacgctccggccgtcgcgagcctcggcacctcgag		72

Qy	913	tgctggctggggctgctctctcttcttaccaccgcgggggcaaggccctggaggagccagca	971
Db	899	TGCTGGCTGGGGCTGGTC--TCCTGTATACACGCCCGGGGCAAGG-CCTGGAGGAGCCAG-CA	954
Qy	973	atgatcaaggaggaatgc	991
Db	955	ATGATATCAAGGAGGATGC	973
RESULT	13		
AL576327/c		AL576327	
LOCUS			
16-FEB-2001			
DEFINITION	AL576327 LTI_NFL006_P12 Homo sapiens cDNA clone CS0DI073YP08 3	978 bp	linear
ACCESSION	AL576327		
VERSION	AL576327.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 978)		
TITLE	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
CONTACT	Contact: Genoscope		
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.			
FEATURES	Location/Qualifiers		
source	1..978		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="CS0DI073YP08"			
/clone_lib="LTI_NFL006_P12"			
/tissue_type="placenta"			
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371			
Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com"			
BASE COUNT	219 a 248 c 297 g 204 t 10 others		
ORIGIN			
Query Match	49.4%	Score 896; DB 9; Length 978;	
Best Local Similarity	97.1%	Pred. No. 1.8e-201;	
Matches 943; Conservative	9; Mismatches 15; Indels 4; Gaps		
4;			
Qy	799	aggcccaaatgaggtggcactgcccaatgtaatgacgcctggagtgagcacaggcc	958

Db	80	CTGTGGTAGTGGGAGAGATAGGAATCTTTATTAAACTACATGAANT	21	
Qy	1758	atgtgtgtttt 1768		
Db	20	ATTGTGTAT 10		
RESULT 14				
AL546335				
LOCUS	AL546335	938 bp	mrna	EST
16-FEB-2001				
DEFINITION	AL546335 IT1_NFL006_PL2 Homo sapiens cDNA clone CS0D1031Y109 5 prime, mRNA sequence.			
ACCESSION	AL546335			
VERSION	AL546335.1 GI:12879351			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	JOURNAL Published (2001)			
COMMENT	Contact: Genoscope			
FEATURES	Genoscope - Centre National de Sequencage			
	BP 191 91006 EVRY cedex - France			
	Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr.			
	Location/Qualifiers			
	1. .938			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="CS0D1031Y109"			
	/clone_lib="IT1_NFL006_PL2"			
	/tissue_type="placenta"			
	/note="Vector: pCMVSPORT 6; Site_1: NotI, 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"			
	BASE COUNT 168 a 282 c 205 t 2 others			
ORIGIN				
Query Match	49.4%; Score 895.8; DB 9; Length 938;			
Best Local Similarity	99.0%; Pred. No. 2e-201;			
Matches 930; Conservative	2; Mismatches 4; Indels 3; Gaps 3;			
Qy	10	ctgggtgtcagcggtcggctccgcgcagcgtccggcgtcgccgagcctcggcacctg	69	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2002, 16:13:17 ; Search time 57.91 Seconds
(without alignments)

updates/sec 748.036 Million cell

Title: US-09-902-759-39

Perfect score: 2012

Sequence: 1 MISLPGPLVTLNLLRFLGL.....SRMGAVPMVPAQSGSLV 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneexp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	2012	100.0	390	20	AAV27096 Human viral recept
2	2012	100.0	390	20	AAV13351 Amino acid sequenc
3	2012	100.0	390	20	AAV05286 EGF-like homologue
4	2012	100.0	390	21	AAV88574 Human PRO246 amino
5	2012	100.0	390	21	AAV94999 Human secreted pro
6	2012	100.0	390	22	AAU12340 Human PRO246 polyp
7	2012	100.0	390	22	AAV88358 Human membrane or
8	2012	100.0	390	22	AAV88599 PRO246. Homo sapi
9	2012	100.0	390	22	AAV81207 Amino acid sequenc
10	2012	100.0	390	22	AAV80219 Human PRO246 prote
11	2012	100.0	390	22	AAV83082 Human angiogenesis
12	2004	99.6	390	22	AAE06610 Human protein havi
13	2004	99.6	390	22	AAV90818 Human shear stress
14	2003	99.6	389	21	AAV76303 Fragment of human
15	1738.5	86.4	370	22	AAV65832 Human INTERCEPT 25
16	1736.5	86.3	370	22	AAV65906 Human secreted pro
17	1734.5	86.2	370	22	AAV65904 Human secreted pro
18	1734.5	86.2	370	22	AAV65905 Human secreted pro
19	1734.5	86.2	370	22	AAV65907 Human secreted pro
20	1601.5	79.6	341	22	AAV65833 Murine mature INTE
21	1581	78.6	321	22	ABV11937 Human viral recept
22	1581	78.6	321	22	AAV40551 Human polypeptide
23	1579	78.5	325	21	AAV95024 Human clone vcs1_1
24	1399	69.4	394	22	AAV65910 Murine secreted pr
25	1397	69.4	394	22	AAV65840 Murine INTERCEPT 2
26	1397	69.4	394	22	AAV65908 Murine secreted pr
27	1394	69.3	394	22	AAV65911 Murine secreted pr
28	1393	69.2	394	22	AAV65909 Murine secreted pr
29	1331	66.2	365	22	AAV65841 Murine mature INTE
30	1232	61.2	246	22	AAV65835 Murine INTERCEPT 2
31	1207	60.0	237	21	AAV76152 Human secreted pro
32	1107	55.0	220	22	AAV65872 Human INTERCEPT 25
33	1095	54.4	217	22	AAV65871 Human INTERCEPT 25
34	1027	51.0	206	22	AAV65870 Human secreted pro
35	965.5	48.0	212	20	AAV25748 Human secreted pro
36	890	44.2	177	22	AAV65867 Murine INTERCEPT 25
37	833.5	41.4	249	22	AAV65843 Murine INTERCEPT 2
38	833	41.4	172	22	AAV38765 Human polypeptide
39	577	28.7	182	22	AAE05354 Mouse 10.3 kDa pro
40	467.5	23.2	120	22	AAV65846 Murine INTERCEPT 2
41	436	21.7	127	21	AAV95025 Human clone vcs1_1
42	400.5	19.9	99	22	AAV65837 Murine INTERCEPT 2
43	400	19.9	80	22	AAV65838 Murine INTERCEPT 2
44	367.5	18.3	426	22	ABV10359 Human cDNA SEQ ID
45	361.5	18.0	376	19	AAV57213 Mouse coxsackievir

ALIGNMENTS

RESULT 1
AAV27096
ID AAV27096 standard; Protein; 390 AA.

PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
PS Claim 12; Fig 17; 320pp; English.
XX
XX
CC AAV1344-403 represent secreted and transmembrane human proteins.
CC The cDNA sequences are obtained from cDNA libraries, prepared from
CC fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
CC The encoded polypeptides have specific uses based on their homology to
CC known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
CC associated with the preservation and maintenance of gastrointestinal
CC mucosa and the repair of acute and chronic mucosal lesions
CC (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
CC ulceration and congenital microvillus atrophy), skin diseases associated
CC with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
CC cancers such as lung squamous cell carcinoma of the vulva and gliomas),
CC potent effects on cell growth and development, diseases related to
CC growth or survival of nerve cells including Parkinson's disease,
CC Alzheimer's disease, AIDS, neuropathies or cancer. PRO265 can be used as
CC for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
CC of Usher Syndrome or Atrophica areata; PRO269 can be used as an
CC anti-thrombotic agent; PRO287 polypeptides and portions may have
CC therapeutic applications in wound healing and tissue repair; PRO317 can
CC be used for treating problems of the kidney, uterus, endometrium, blood
CC vessels, or related tissue, e.g. in the heart of genital tract.
XX
SQ Sequence 390 AA;
Query Match 100.0%; Score 2012; DB 20; Length 390;
Best Local Similarity 100.0%; Pred. No. 5.8e-143;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MISLPGPLVTLNLRFLPLGLSALAPPSPRAQLQLHLPAHLQAVGGGVVLPWYTLHGVEV 60
DB 1 MISLPGPLVTLNLRFLPLGLSALAPPSPRAQLQLHLPAHLQAVGGGVVLPWYTLHGVEV 60
QY 61 SSSQPEWPFVFMFFKQKEKEDVLSYNGVTTSGKSVLSVSNPGRNLSRLGLQEKD 120
DB 61 SSSQPEWPFVFMFFKQKEKEDVLSYNGVTTSGKSVLSVSNPGRNLSRLGLQEKD 120
QY 121 SGFYSCSVNVDPQKSGRHSKITLEMLVLPAPPSCNLOGVPHGVNVLSCQSPRSK 180
DB 121 SGFYSCSVNVDPQKSGRHSKITLEMLVLPAPPSCNLOGVPHGVNVLSCQSPRSK 180
QY 181 PAVOVNDROLPSFOTFPAPALDVRIGSLTSLNLSMAGVYVCKKARHNVGTACNVLTLE 240
DB 181 PAVOVNDROLPSFOTFPAPALDVRIGSLTSLNLSMAGVYVCKKARHNVGTACNVLTLE 240
QY 241 VSTGPGAAVAGAVGTVLVGLIAGLVLYHRGKALEEPDANKEDAIAPRLPMPKS 300
DB 241 VSTGPGAAVAGAVGTVLVGLIAGLVLYHRGKALEEPDANKEDAIAPRLPMPKS 300
QY 301 SDTISKNGTSLSSVSARALRPPHPPRGCAUTPTPSLSQALPSRLPTTGAHPQPISF 360
DB 301 SDTISKNGTSLSSVSARALRPPHPPRGCAUTPTPSLSQALPSRLPTTGAHPQPISF 360

PR 25-NOV-1997; 97US-0065840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065893.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.
PA (GETH) GENENTECH INC.
PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
XX WPI; 1999-229533/19.
DR N-PSDB; AAX52221.
XX

CC may be conjugated to a toxin, chemotherapeutic agent or radioisotope.
 CC Genes expressing (I), many of which are growth factor homologues, are
 CC overexpressed in some cases of cancer.
 XX
 SQ Sequence 390 AA;

Query Match 100.0%; Score 2012; DB 20; Length 390;
 Best Local Similarity 100.0%; Pred. No. 5.8e-143;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISLPGPLVTNLLRFLTGLSALAPPSPAOLQLHLPAERLQAVESGVEVLPAAWTLTGGV 60
 |||||||
 DB 1 mislpgplvtnllrflfiglsalappsqalqihpanriqavessvlpawtllngsv 60
 |||||||

QY 61 SSSQWEVPMVPMFFKQEKEDQVLSYINGVTTSPKGVSLVSNPSENLSRLSLGLQEKD 120
 |||||||
 DB 61 ssqpwepvmvpmffkqekedqvlsyngvttspkgsylvsnpsemlslrlslglqekd 120
 |||||||

QY 121 SGFYSQVNTQDKQKSRGHSIKTLEMLVLPAPPSCRLQGVPHVGNVTLSCQSPRSK 180
 |||||||
 DB 121 sgpyacsvnvqdkqksgkgahtlelnvlpappscrlqgvphvganvtlscqgsprsk 180
 |||||||

QY 181 PAVQVQMDRLPSEFTFFPALDVRIGSLSTNLSSMAGVYVCKAHNEVGTACNVTL 240
 |||||||
 DB 181 pavqvmdrlpseftffapaldvirgslstnlssmagvvyckahnevgtaqcnvtle 240
 |||||||

QY 241 VSTGGAANVAGNVGTGLGLLGLLXHRGKALEEPAANDIKEDAIAPRLPWPKS 300
 |||||||
 DB 241 vstggaanvagnvgtglglglxhrgkaleepaandikedaiaprlpwpks 300
 |||||||

QY 301 SDTISKNTLSSVTGARALRPPHPPRCALTTPSLSSQALPSRLPTTDCAHPOQISP 360
 |||||||
 DB 301 sdtiskntlssvtagaralrpphpprcalttptslssqalpsrlpttdcahpqois 360
 |||||||

QY 361 IPGVSSSLSRMGAVPMVPAQSQAGSLV 390
 |||||||
 DB 361 ipgvssslsrmgavpmvpaqsqagslv 390
 |||||||

RESULT 4
 AAY88574
 ID AAY88574 standard; Protein; 390 AA.
 XX
 XX AAY88574;
 XX AC
 XX DT 09-AUG-2000 (first entry)
 XX DE Human PRO246 amino acid sequence.
 XX KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261;
 PRO246;
 KW PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;
 KW cell growth; proliferation; cell surface virus receptor; ADPRT;
 KW antibody dependent enzyme mediated prodrug therapy.
 XX

DB 361 ipgvssslsrmgavpmvpaqsqagslv 390
 |||||||

RESULT 3
 AAY05286
 ID AAY05286 standard; Protein; 390 AA.
 XX
 AC AAY05286;
 XX
 DT 22-JUN-1999 (first entry)
 XX
 DE EGF-like homologue PRO246.
 XX
 KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261;
 PRO246;
 KW EGF-2; inhibitor; tumour growth; cancer; EGF-like homologue;
 KW FGF-8 homologue.
 XX
 OS Homo sapiens.
 XX
 PN W09914327-A2.
 XX
 PD 25-MAR-1999.
 XX
 PF 10-SEP-1998; 98WO-US18824.
 XX
 PR 25-NOV-1997; 97US-0065840.
 PR 17-SEP-1997; 97US-0059114.
 PR 17-SEP-1997; 97US-0059117.
 PR 18-SEP-1997; 97US-0059263.
 PR 15-OCT-1997; 97US-0062125.
 PR 17-OCT-1997; 97US-0062285.
 PR 17-OCT-1997; 97US-0062287.
 PR 24-OCT-1997; 97US-0062816.
 PR 29-OCT-1997; 97US-0063704.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Botstein D, Goddard A, Gurney A, Hillan K, Lawrence DA;
 PI Roy M, Wood WI;
 XX
 DR WPI; 1999-229532/19.
 DR N-PSDB; AAY28436.
 XX
 FT Antibodies against specific proteins overexpressed in tumours
 XX
 PS Example 1; Fig 27; 130pp; English.
 XX
 CC This sequence represents the EGF-like homologue PRO246.
 CC The invention relates to antibodies (Ab) that bind to any of the
 CC polypeptides (I) designated PRO187; PRO533; PRO214; PRO240; PRO211;
 CC PRO230; PRO261; PRO246 or EGF-2. The Ab, or other agents that inhibit
 CC expression and/or activity of (I) are used: (i) to inhibit growth of
 CC tumours; and (ii) as diagnostic/prognostic reagents for detection or
 CC quantification of (I) in cells or tissues, by standard immunoassays,
 with

OS	Homo sapiens.
PX	WO200015666-A2.
XX	
XX	PD 23-MAR-2000.
XX	
FF	08-SEP-1999; 99WO-US20594.
PR	10-SEP-1998; 98US-0099803.
PR	10-SEP-1998; 98WO-US18824.
XX	(GETH) GENENTECH INC.
PA	Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Botstein D;
PI	WFI; 2000-271386/23.
DR	N-PDSB; AAA30052.
XX	
PT	New isolated antibodies which bind to specific polypeptides used for
PT	diagnosis and treatment of neoplastic cell growth and proliferation -
XX	
PS	Example 8; Fig 16; 200pp; English.
XX	This sequence represents a human PRO246 amino acid sequence. PRO246 is
CC	probably a cell surface virus receptor. The invention relates to
CC	isolated
CC	antibodies which bind to a polypeptide. The "PRO" polypeptides are
CC	encoded by genes which are over expressed in the genome of tumour cells.
CC	Vectors and host cells comprising the nucleic acid encoding the
CC	antibodies are used in the production of the antibodies. The antibodies
CC	and nucleic acids encoding them are used for diagnosing a tumour in a
CC	mammal. The antibodies are used for inhibiting the growth of tumour
cells	
CC	and identifying compounds that inhibit a biological or immunological
CC	activity of and/or expression of a PRO187, PRO533, PRO214, PRO240,
CC	PRO211, PRO230, PRO261, PRO246 or PRO317 polypeptide. The antibody can
be	used in antibody dependent enzyme mediated prodrug therapy (ADEPT) by
a	conjugating the antibody to a prodrug-activating enzyme which converts
CC	prodrug to an anti-cancer drug. The antibodies can be fluorescently
fluorimetry	labelled and monitored by light microscopy, flow cytometry or
CC	for diagnosis and prognosis of tumours.
XX	
SQ	Sequence 390 AA;
	Query Match 100.0%; Score 1012; DB 21; Length 390;
	Best Local Similarity 100.0%; Pred. No. 5.8e-143;
	Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps
0;	
QY	1 MISLPGPLVNLIRFLFGLSALAPPSPRAQLQHLPANRLQAVGGGVVLPAWYTLHGVEV 60
DB	1 mislpgplvnlrlrflfglsalappsqqlqihlpnarlqavgggervvpawylhgvev 60
OY	61 SSSOWEPVPVWHPFKOKEKEDOVLSYINGVTTSKPGVSLVYSWPSRSLISLRLEGLOEKD 120

АБВГДЕЖЗИЙКАЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ

241 VSTGPGAAVWAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIPTLTPWPKS 300

360
301 edciakngtisvtsaralrphgpprfgaltptpalaeqalpprriptdgahqppisp 360

Ddb

361 IPGVSSSGLSRMGAVPVMVPAQSQAGSLV 390

QY

361 ipgvysssgslrmgavpvmvpaqsqagslv 390

Ddb

361 ipgvysssgslrmgavpvmvpaqsqagslv 390

RESULT	6	
AAU12340		
ID	AAU12340	standard; Protein; 390 AA.
XX		
AC	AAU12340;	
XX		
DT	24-OCT-2001	(first entry)
XX		
DE	Human PRO246	polypeptide sequence.
XX		
KW	Human secretory and transmembrane;	PRO; mammalian; cancer; lung;
KW	Breast; prostate;	cervical; tumour necrosis factor-alpha; TNF-alpha;
KW	cartilage; ear;	proliferation; glucose; free fatty acid; skeletal
KW	muscle;	
KW	adipocyte; A-peptide;	factor VIIA; gene therapy.

01-DEC-2000: 2000WO-US32678.

PR	09-DEC-1999;	99US-0170262.
PR	16-DEC-1999;	99WO-US30095.
PR	20-DEC-1999;	99WO-US30911.

PR	30-DEC-1999;	99WO-US31243.
PR	06-JAN-2000;	2000WO-US00277.
PR	06-JAN-2000;	2000WO-US00376.

PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1 mislpgplvtlnlrrlrltliglsalappetradiqlnlpamrlqavgevevlpawyltngav bu

61 SSSQPEWVFPVMMFFKQKEKEDQVLSYINGVTTSKPGVSLVSYMPNRNLSRLRGLEQKD 120
 61 sssqpewevpvmvffkqekedqvlsyngvttstkpgvalsymparnlslrlgileqkd 120
 121 SGFYSCVNVQDKQKSGRHSIKTLELVNLYVPAPPSCELOQVPHVGNVTLSCQSPRSK 180
 121 sgpyscvnnvqdkgkgrghsiktelnlvlyvpappscrlqgvphvganvtlscqspresk 180
 181 PAVQVNDROLPSQTFPPALDVRGSLVTLNLSMAGVYVCKKAHNEVGTACQNVTLLE 240
 181 pavqvgndrlpqtfppaladvrgslvtlnlsmagvyvckkahnnevgtaqcnvtlle 240
 241 VSTGPGAAVAGAVGTVLGLGLAGLVLLVHRGKALEEPANDIKEDAIAPRTLPMWPKS 300
 241 vstgpgaaavagavgtvlgllaglvllvhrhgkaleepandikedaiaprtlpmpks 300
 301 SDTISXNGTSSVTSARALRPPHPPRPGALTPPTSLSSQALPSPRLPTTGDHAPQIPSP 360
 301 sdtkngtssvtsaralrpphprpgaltptpslssqalpsprlpttdghapqipsp 360
 361 IPGGVSSSGLSRMGAVFVMPVPAQSQAGSLV 390
 361 ipgvgssglsrmgavfvmvpqdsqagslv 390

RESULT 7
 AAB88358
 ID AAB88358 standard; Protein; 390 AA.
 XX AAB88358;
 AC AAB88358;
 DT 23-MAY-2001 (first entry)
 DE Human membrane or secretory protein clone PSEC0086.
 KW Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW rheumatoid arthritis; diabetes.
 OS Homo sapiens.
 XX EP1067182-A2.
 XX 10-JAN-2001.
 XX 07-JUL-2000; 2000EP-0114030.
 XX 08-JUL-1999; 99JP-0194179.
 XX 11-JAN-2000; 2000JP-0118775.
 XX 02-MAY-2000; 2000JP-0183766.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 XX WPI, 2001-093989/11.
 XX N-PSDB; AAF93785.

24-FEB-2000; 2000HO-US05004.
 PR 01-MAR-2000; 2000HO-US05601.
 PR 20-MAR-2000; 2000HO-US07377.
 PR 21-MAR-2000; 2000HO-US07532.
 PR 30-MAR-2000; 2000HO-US08439.
 PR 17-MAY-2000; 2000HO-US13705.
 PR 22-MAY-2000; 2000HO-US14042.
 PR 30-MAY-2000; 2000HO-US14941.
 PR 02-JUN-2000; 2000HO-US15264.
 PR 10-NOV-2000; 2000HO-US30873.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood NI, Zhang Z;
 XX WPI; 2001-408281/43.
 DR N-PSDB; AAS21412.
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX Claim 12; Fig 338; 813pp; English.
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMNCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX Sequence 390 AA;
 SQ
 Query Match 100.0%; Score 2012; DB 22; Length 390;
 Best Local Similarity 100.0%; Pred. No. 5.8e-143;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps
 0;
 QY 1 MISLPGPLVTNLLRPLFLGLSLNAPPSRAOLQLHPANLSAVGEGEVLPATWTLHGEV 60
 |||||

Qy	301	SDATISKNGTSSVTSARALRPPHCPPPRPGALTPPTSSLSQALPSRLFTTDCGHPQPTSP	360
Db	301	atdkisngkngtlssvtsaralrpphchpprpgaltpptsslsqalpsrlfttcdgahpqpisp	360
Qy	361	IPGCVSSSGLSRMGAVFVMPVPAQSQAGSLV	390
Db	361	lpgsvsssglsrmgavfvmvmpaqagagslv	390
RESULT	8		
AAB68599			
ID	AAB68599	standard; Protein, 390 AA.	
AC	AAB68599		
DT	27-APR-2001	(first entry)	
DE	PRO246.		
KW	Cytostatic; PRO protein; tumour; cancer.		
OS	Homo sapiens.		
FN	WO200105836-A1.		
PD	25-JAN-2001.		
PF	20-DEC-1999; 99WO-US30999.		
PR	20-JUL-1999; 99US-0144758.		
PR	26-JUL-1999; 99US-0145698.		
PR	08-SEP-1999; 99WO-US20594.		
PR	13-SEP-1999; 99WO-US20944.		
PR	15-SEP-1999; 99WO-US21090.		
PR	05-OCT-1999; 99WO-US23089.		
PR	29-NOV-1999; 99WO-US28214.		
PR	30-NOV-1999; 99WO-US28313.		
PR	02-DEC-1999; 99WO-US28564.		
PA	(GETH) GENENTECH INC.		
PI	Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;		
XX	WFI; 2001-091968/10.		
DR	N-PSDB; AAF60372.		
PT	New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,		
PT	useful for diagnosing and treating cancers -		
PS	Claim 61; Fig 16; 196pp; English.		
XX	The present invention relates to PRO proteins and coding sequences. The		
CC	present sequence is one such PRO protein. It was found that the PRO		
genes			
CC	are amplified in the genome of tumour cells. The gene amplification is		

CC expected to be associated with the overexpression of the gene product
 and
 CC contributes to tumorigenesis. Therefore, antagonists of PRO proteins
 are
 CC useful for the treatment of benign or malignant tumours, leukaemias,
 CC lymphoid malignancies and other disorders such as neuronal, glial,
 CC astrocytal, hypothalamic, glandular, epithelial, inflammatory and
 CC immunologic disorders.
 XX
 SQ Sequence 390 AA;

Query Match 100.0%; Score 2012; DB 22; Length 390;
 Best Local Similarity 100.0%; Pred. No. 5.8e-143;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISLPGPLVTLNLRFLFGLSALAPPRAQIQHLHPANLQAVGGSEVVLPAWTLHGVE 60
 DB 1 mislpgplvtlnlrflfiglslalappaqqlhlpanrlqavggsevvlpawtylhgev 60
 QY 61 SSSQPEVFFVFWFFKQEKEDQVLYNGVTTSKFGVSLVYSPRNLSLRLGLEQED 120
 DB 61 sssqpevffvfwffkqekedqvlvngvttstkgvslvysprnlslrlgleqkd 120
 QY 121 SGYSCSNVNDVQKSGSKHKTLELVLPAPPSCELQGVHVCANVTLSQSPRSK 180
 DB 121 sgyscsnvndvqksgskhktlelvlpappscelqgvhvcantlscqsprk 180
 QY 181 PAVQYQDRQLPQSFQFPAPALDVIRGSLTINLSMAGVVCVKAHNVGPAQCNVTL 240
 DB 181 pavqyqdrqlpsqfppapaldvirgslstlnlsmagvvcvkaahnvvgpaqcnvtle 240
 QY 241 VSTPGAAVAVGAVVGLVGLLGLVLLVYHRRGKALEEPANDIKEDAIAPRTLPHPKS 300
 DB 241 vstpgaaavvagavvgtlvglgllagvllyhrrgkaleepandikedaiaprtlpwks 300
 QY 301 SDTISKNGTSLSVTSARALRPHGPFPQGNLTPTPSLSQALPGRPLPTDGHAPQISF 360
 DB 301 sdtkngtslsvtsaralrphgfpqgnltptpslsqalpgrplptdghapqisfp 360
 QY 361 IPGGVSSSLSEKGNVPMVPAQSQAGSLV 390
 DB 361 ipggvssslsekgnvpmvpaaqagalsv 390

RESULT 9

AAB31207
 ID AAB31207 standard; Protein; 390 AA.

XX AAB31207;
 XX 20-APR-2001 (first entry)
 XX Amino acid sequence of human polypeptide PRO246.
 DE Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
 XX PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361;

PRO1308;
 KW PRO1183; PRO1272; PRO1419; PRO4999; PRO1710; PRO248; PRO353; PRO1318;
 KW PRO1600; PRO3940; PRO333; PRO301; PRO187; PRO337; PRO411; PRO4356;
 KW PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;
 KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy.

OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT Location/Qualifiers
 FT 1..29
 FT /note= "signal peptide"
 FT Modified-site 90..96
 FT /note= "N-myristoylation site"
 FT Modified-site 108..112
 FT /note= "N-glycosylation site"
 FT Modified-site 167..173
 FT /note= "N-myristoylation site"
 FT Modified-site 169..173
 FT /note= "N-glycosylation site"
 FT Modified-site 213..217
 FT /note= "N-glycosylation site"
 FT Modified-site 220..226
 FT /note= "N-myristoylation site"
 FT Modified-site 231..237
 FT /note= "N-myristoylation site"
 FT Modified-site 236..240
 FT /note= "N-glycosylation site"
 FT Domain 245..267
 FT /note= "transmembrane protein"
 FT Modified-site 252..258
 FT /note= "N-myristoylation site"
 FT Modified-site 256..262
 FT /note= "N-myristoylation site"
 FT Modified-site 262..268
 FT /note= "N-myristoylation site"
 FT Modified-site 307..311
 FT /note= "N-glycosylation site"
 FT Modified-site 308..314
 FT /note= "N-myristoylation site"
 FT Modified-site 363..369
 FT /note= "N-myristoylation site"
 FT Modified-site 364..370
 FT /note= "N-myristoylation site"

WO200077037-A2.

XX 21-DEC-2000.
 XX 22-MAY-2000; 2000WO-US14042.
 XX 15-JUN-1999; 99US-0139695.
 PR 20-JUL-1999; 99US-0145070.
 PR 26-JUL-1999; 99US-0145698.
 PR 17-AUG-1999; 99US-0149396.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.

07-DEC-1999; 99US-0169495.
 PR 07-DEC-1999; 99US-0169495.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 15-MAY-2000; 2000WO-US13358.
 PR 17-MAY-2000; 2000WO-US13705.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Botstein DA, Deanovsers L, Eaton DL;
 PI Ferrara N, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
 PI Godowski PJ, Gurney AJ, Kijavini IJ, Mather JP, Napier MA, Pan J;
 PI Padoni WF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PW;
 PI Wood WI, Zhang Z;
 XX
 DR WFI; 2001-050091/06.
 DR N-PSDB; AAC87040.
 XX
 PS Claim 12; Fig 58; 24pp; English.
 CC
 CC The present sequence represents a human secreted and transmembrane
 CC polypeptide. The specification describes human polypeptides, designated
 CC PRO196, PRO444, PRO183, PRO210, PRO215, PRO217, PRO242, PRO288,
 CC PRO365, PRO1361, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
 CC PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
 CC PRO317, PRO1411, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
 CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells
 CC can be modulated with agents that bind to these polypeptides, resulting
 CC in the death of the cells. The polynucleotides encoding these
 CC polypeptides are useful in the recombinant production of the
 CC polypeptides, as a hybridisation probe to screen libraries to isolate
 CC homologous sequences, or to map the gene. They may also be used for
 CC analysing genetic disorders, and to produce transgenic animals which are
 CC useful for the development and screening of therapeutically useful
 CC reagents. The polynucleotides can also be used in gene therapy e.g. to
 CC replace a defective gene.
 XX
 SQ Sequence 390 AA;
 Query Match 100.0%; Score 2012; DB 22; Length 390;
 Best Local Similarity 100.0%; Pred. No. 5.8e-143;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps
 0;
 QY 1 MISLPGPLVTLRLFLGLSALAPPSRAQLQLHPANRLQAVEGGEVVLPAWYTLRGEV 60

QY 61 SSSQWEVPVWPFKKKEKEDVLSYINGVTTTSKPGVSLVYSWPSRNLRLRLEGLQEKD 120
 DB 61 SSSQWEVPVWPFKKKEKEDVLSYINGVTTTSKPGVSLVYSWPSRNLRLRLEGLQEKD 120
 QY 121 SGPYSCSVNVQDKQKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGVANVTLSCQSPRSK 180
 DB 121 SGPYSCSVNVQDKQKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGVANVTLSCQSPRSK 180
 QY 181 PAVQYQMDRLQSFQTFPAPLDVIRGSLSTNLSSMAGVYVCKAHNEVGTAQCKNVTL 240
 DB 181 PAVQYQMDRLQSFQTFPAPLDVIRGSLSTNLSSMAGVYVCKAHNEVGTAQCKNVTL 240
 QY 241 VSTGGAAVVAGAVVGTVLGVLGSLAGLVLLYHRRKALKEEPANDIKEDAIAPRTLWPWKS 300
 DB 241 VSTGGAAVVAGAVVGTVLGVLGSLAGLVLLYHRRKALKEEPANDIKEDAIAPRTLWPWKS 300
 QY 301 STTISKNGTILSSVTSARALPPHGPFRGALTPTSLSSQALSPRLPTTDGAPPOISP 360
 DB 301 STTISKNGTILSSVTSARALPPHGPFRGALTPTSLSSQALSPRLPTTDGAPPOISP 360
 QY 361 IPGGVSSSLGRMGAVPMVVPASQAGSLV 390
 DB 361 IPGGVSSSLGRMGAVPMVVPASQAGSLV 390
 RESULT 10
 AAB80219
 ID AAB80219 standard; Protein; 390 AA.
 XX
 AC AAB80219;
 DT 24-APR-2001 (first entry)
 XX
 DE Human PRO246 protein.
 KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
 KW antiparkinsonian neurotropic; neuroprotective; vulnary; cardiant;
 KW antiangiosgenic; vasotropic; antiasthmatic; antirneumatic; cancer;
 KW antithratic; antinfertility; antidiabetic; antiviral; diabetes;
 KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
 KW ischaemia; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200104311-A1.
 XX
 PD 18-JAN-2001.
 XX
 PF 22-FEB-2000; 2000WO-US04414.
 XX
 PR 07-JUL-1999; 99US-0143048.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.

PR	15-SEP-1999;	99WO-US21090.	
PR	15-SEP-1999;	99WO-US21547.	
PR	05-OCT-1999;	99WO-US23089.	
PR	29-NOV-1999;	99WO-US28214.	
PR	30-NOV-1999;	99WO-US28313.	
PR	16-DEC-1999;	99WO-US30095.	
PR	20-DEC-1999;	99WO-US30911.	
PR	20-DEC-1999;	99WO-US30999.	
PR	05-JAN-2000;	99WO-US00219.	
XX			
PA	(GETH)	GENENTECH INC.	
XX			
PI	Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;		
PI	Pilvaroff E, Fong S, Gao W, Gerber H, Gerritsen MB, Goddard A;		
PI	Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kijavini LJ;		
PI	Maccher JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;		
PI	Williams PM, Wood WI;		
XX			
DR	WPI; 2001-081051/09..		
DR	N-PSDB; AAF72379.		
XX			
PT	Sixty one nucleic acids encoding PRO polypeptides which are useful in		
PT	the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung		
PT	squamous cell carcinoma) and neurodegenerative diseases (e.g.		
PT	Alzheimer's disease) -		
XX			
PS	Claim 1; Fig 17; 393pp; English.		
XX			
CC	The present sequence is one of sixty one novel secreted and		
CC	transmembrane PRO polypeptides. The PRO polypeptides are		
CC	useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung		
CC	squamous cell carcinoma), gastrointestinal disorders (e.g.		
CC	enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,		
CC	Parkinson's disease), wound repair, cardiovascular disorders (e.g.		
CC	endometrial bleeding angiogenesis, ischaemias such as coronary		
CC	ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,		
CC	rheumatoid arthritis, multiple sclerosis), infertility, AIDS and		
CC	diabetes and retinal disorders such as retinitis pigmentosum.		
CC	The PRO nucleic acids have applications in molecular biology, including		
CC	use as hybridization probes, and in chromosome and gene mapping.		
XX			
SQ	Sequence 390 AA;		
	Query Match 100.0%; Score 2012; DB 22; Length 390;		
	Best Local Similarity 100.0%; Pred. No. 5.8e-143;		
	Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps		
0;			
QY	1 MISLPGLVTLNLRFLFGLSALAPPSRAQLQLHPANLRQAVRGGEVVLPAWYTLHGVS 60		
Db	1 mlslpplvtnllrfiflglsalapparaqlqlhpanrlqavgeevlpawtytlhgev 60		
QY	61 SSSQPEVFFVFMFFKQEKEDQVLYINGVTTSKPGVSLVYSPNSLRLSLGLQEKD 120		
Db	61 ssqpevffvfmffkqekedqvlyeingvttakgvalvyemparnislrlslgkqd 120		
QY	121 SGYSCSVNVODKQKSRGHSIKTLELVLPAPPSCRLQGVPHGVANVTUSCQSPRK 180		

Db	121	sgpyscsvnvodkqkargkshkltlelvlpappscrlqgvphgvnvanvscsqprsk	180
QY	181	pavqvqwdrlpfsfotffafpaldvirsrltlnlsssmagvtyckahrnevgtacnvtile	240
Db	181	pavqvqwdrlpfsfotffafpaldvirsrltlnlssamagvtyckahnevgtacnvtile	240
QY	241	vtstgpaavagavvgtlvgilagvllyhrrgkalesepandikeaiaiprtlpwfk	300
Db	241	vtstgpaavagavvgtlvgilagvllyhrrgkalesepandikeaiaiprtlpwfk	300
QY	301	sditiskngtltssvtsaralpprhgpprcaltpptslssqalspelttdgahropisp	360
Db	301	sditiskngtltssvtsaralpprhgpprcaltpptslssqalspelttdgahpapiap	360
QY	361	ipggyssgslsrmgavpvmvpaqsgslv	390
Db	361	ipggyssgslsrmgavpvmvpaqsgslv	390
RESULT 11			
ID	AA053082	standard; Protein; 390 AA.	
XX			
AC	AA053082;		
XX			
DT	28-FEB-2001	(first entry)	
XX			
DE		Human angiogenesis-associated protein PRO246, SEQ ID NO:96.	
XX			
KW		Human; angiogenesis-associated protein; PRO; endothelial cell growth;	
KW		cardiac hypertrophy; cardiovascular disorder; endothelial disorder;	
KW		angiogenic disorder; atherosclerosis; osteoporosis; hypertension;	
KW		myocardial infarction; diabetic retinopathy; rheumatoid arthritis;	
KW		Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;	
KW		Alzheimer's disease; Huntington's disease; stroke; drug screening;	
KW		gene therapy; transgenic animal.	
XX			
OS		Homo sapiens.	
XX			
PN		WO200053753-A2.	
XX			
PD		14-SEP-2000.	
XX			
XX		05-JAN-2000; 2000WO-US00219.	
PF			
XX			
PR		08-MAR-1999; 99WO-US05028.	
PR		12-MAR-1999; 99US-0123957.	
PR		14-MAY-1999; 99US-0134287.	
PR		02-JUN-1999; 99WO-US12252.	
PR		23-JUN-1999; 99US-0141037.	
PR		20-JUL-1999; 99US-0144758.	
PR		26-JUL-1999; 99US-0145698.	
PR		01-SEP-1999; 99WO-US20111.	
PR		08-SEP-1999; 99WO-US20594.	
PR		15-SEP-1999; 99WO-US21090.	
PR		15-SEP-1999; 99WO-US21547.	
PR		05-OCT-1999; 99WO-US23089.	

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;

PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;

PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;

XX MPI; 2001-090793/10.

DR N-PSDB; AAC97441.

XX New isolated nucleic acid for producing a PRO polypeptide, analyzing

PT genetic disorders and treating cardiovascular, endothelial or

PT angiogenic disorders, such as atherosclerosis, wounds or cancer -

XX Claim 69; Fig 38; 293pp; English.

XX The invention relates to novel human angiogenesis-associated proteins

CC designated PRO proteins (AA053064-B53097), and to nucleic acids encoding

CC PRO proteins. The invention also relates to vectors and host cells

CC comprising a PRO nucleic acid, the recombinant production of a PRO

CC protein, PRO antibodies specific for a PRO protein, fusion proteins

CC comprising a PRO protein, agonists or antagonists of a PRO protein, and

CC compounds which inhibit the expression of a PRO gene. The invention

CC additionally encompasses methods of identifying modulators of PRO

CC expression or activity; diagnosing a cardiovascular, endothelial or

CC angiogenic disorder, or a susceptibility to such a disorder by detecting

CC mutations in a PRO gene, or the expression level of a PRO gene within a

CC particular tissue; treating a cardiovascular, endothelial or angiogenic

CC disorder via the administration of a PRO protein, PRO nucleic acid, or

CC PRO agonist or antagonist; a retroviral gene therapy vector comprising

a PRO nucleic acid; and methods of inhibiting or stimulating endothelial

CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the

CC administration of a PRO protein, or an agonist or antagonist thereof.

CC PRO nucleic acids, PRO proteins, antibodies against PRO proteins, PRO

CC agonists and PRO antagonists may be used as therapeutic agents to treat

CC cardiovascular, endothelial or angiogenic disorders, such as

CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,

CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,

CC endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's

CC disease, or stroke. PRO nucleic acids are additionally useful in the

CC recombinant production of PRO proteins, as hybridisation probes to

CC screen libraries to isolate cDNAs with sequence identity to PRO

proteins,

CC to map genes encoding PRO proteins, to analyse genetic disorders, and in

CC gene therapy. PRO nucleic acids can also be used to produce transgenic

CC animals useful for the development and screening of potential

CC therapeutic agents. The present sequence represents a PRO protein of the

CC invention.

XX Sequence 390 AA;

Query Match 100.0%; Score 2012; DB 22; Length 390;

QY	1	MISLPGLVTLNLRFLFLGLSALAPPSPRAQLQHLPANRLQAVGEVGVLPANTYILGEV	60
DB	1	mislpvglvtnlrlflglalsalappspraqqlhlpnrlqavgegvlpawylhgev	60
QY	61	SSSQPHEVFPVWFVKQKEDQVLSYINGVTTSKPGVSLVYSPSRLSLRLEGLQEKD	120
DB	61	sssqphevfpvwmvfkkqekedqvlsyngvttskpgvalvysmpsrnlrlrieglqekd	120
QY	121	SGPYSCSVNVQDKGKSRGHSIKTLELVLPAPPSCRLQGVPHVGANVTLSQSPSPSK	180
DB	121	sgpyscsvnvqdkgksgksiktlevlpappscrlqgvphvganvntlsqspspsk	180
QY	181	PAVOYQWDRQLPSQTFPFPALDVIKRSLSLNLSSMAGVYVYKARHVEGTACQNVTLLE	240
DB	181	pavoyqwdrlpsqtfpfpaldvirgslslnlssmagsvvyckabnvegtacqnmvtle	240
QY	241	VSTGFGAANVAGAVVGTILVGLLGLAGLVLLYHRRKALEPANDIKEDAIPTLFWPKS	300
DB	241	vstgfgaanvavvgvtlvgllglaglvlllyhrrgkaleepandikedaiptlfpwks	300
QY	301	SPTISKNGTILSSVTSARALPDPHPGPPRQALPTFSLSSQALSPSLPTTIDGHPQPTSP	360
DB	301	sptiskngtllssvtsaralrpphgprrqaltptpslssqalspsrlpttdgahpqpisp	360
QY	361	IPGVSSSGLSRMGAVVMPVPAQSAGSLV	390
DB	361	ipgvsssglsrmgavvmpvpagsagslv	390
RESULT 12			
AA066610			
ID	AA066610	standard; Protein; 390 AA.	
XX	AC	AA066610;	
XX	DT	25-SEP-2001 (first entry)	
XX	DE	Human protein having hydrophobic domain, HP10801.	
XX	KW	Human; hydrophobic domain; gene therapy; nutritional supplement;	
XX	KW	cell proliferation; immunomodulatory; autoimmune disorder;	
XX	KW	antimicrobial;	
XX	KW	multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;	
XX	KW	haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;	
XX	KW	Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;	
XX	KW	haemostatic; thrombolytic; tumour growth inhibitor; anabolic;	
XX	OS	contraceptive; antiinfertility; antiinflammatory.	
XX	OS	Homo sapiens.	
XX	PH	Key	Location/Qualifiers
XX	FT	Peptide	1..390
XX	FT	Protein	/label= signal_peptide
XX	FT	Protein	31..390

PI Kuga T, Sekine S, Nakamura Y, Sugano S;
 XX WPI; 2001-266308/27.
 DR N-PSDB; AAH02949.
 XX
 XX DNA sequences, proteins encoded by them and antibodies against them
 PT useful in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis -
 XX
 XX Claim 35; Page 599-601; 678pp; Japanese.
 XX
 XX The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension.
 XX
 XX Sequence 390 AA;
 SQ
 Query Match 99.6%; Score 2004; DB 22; Length 390;
 Best Local Similarity 99.7%; Pred. No. 2.3e-142;
 Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps
 0;
 Qy 1 MISLPGSLVTLNRLPLGLSLAPPSPRQAQLQHLPLNRLQAVGEGVPLPWTYLHGEV 60
 Db 1 mislpgslvtnllrlfllglsalappraqglhlplnrlqavgegvlpawtylhgev 60
 Qy 61 SSSQFPEVFWWFKKKEKEDVLVINGVTTSKGVSLVYSPSRNLSLRLEGLOEKD 120
 Db 61 sssqfpevfwvmfkkkedqvlvyingvttskgvslvyspsrnlslrleglgekd 120
 Qy 121 SGPTSCSVNVQDKGSRGHSIKLELNLVPPAPPSCRLQGVPHVGVNATLSCQSPRSK 180
 Db 121 sgptscsvnvqdkgkgrghsiktelnlvppappscrlqgvphvgvntlscqspresk 180
 Qy 181 PAVOYQDRQLPSFTFPALDVIKSLNLTSSSMAGVYCKANVEGTACQNVILE 240
 Db 181 pavoyqdrqlpsftfpaldvirgslntlsssmagvycckahvegtacnvile 240
 Qy 241 VSTGFGAANVAGVVGTIVGLGLAGLIVLYHRRGKALEPANDIKEDAIAPRTLPWPKS 300
 Db 241 vstgfgaanvaggvgtivglglaglivlyhrrgkaleepandikedaiaprtlpwps 300
 Qy 301 SDTISKNGTLSSVTSABALPPHGPFPRLGATPTSLSSQALPSRLPTDGAHQPISP 360
 Db 301 sdtiskngtlssvtsaalpphgpfprrlgatptslssqalpsrlptdgahtpisp 360
 Qy 361 IFGVSSSSGLSRMGAVPMVPAQSQAGSLV 390
 Db 361 ifgvsssgslrmgavpmvmpaqsqagslv 390
 RESULT 14
 AAY76303
 ID AAY76303 standard; Protein; 389 AA.

AC AAY76303;
 XX
 XX 23-MAR-2000 (first entry)
 XX
 XX Fragment of human secreted protein encoded by gene 29.
 DE
 XX Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 KW therapy.
 KW Homo sapiens.
 XX MO9958660-AA.
 XX 18-NOV-1999.
 XX
 XX 06-MAY-1999; 99MO-US09847.
 XX 12-MAY-1998; 98US-0085093.
 XX 12-MAY-1998; 98US-0085094.
 XX 12-MAY-1998; 98US-0085105.
 XX 12-MAY-1998; 98US-0085180.
 XX 18-MAY-1998; 98US-0085906.
 XX 18-MAY-1998; 98US-0085920.
 XX 18-MAY-1998; 98US-0085921.
 XX 18-MAY-1998; 98US-0085922.
 XX 18-MAY-1998; 98US-0085923.
 XX 18-MAY-1998; 98US-0085924.
 XX 18-MAY-1998; 98US-0085928.
 XX 18-MAY-1998; 98US-0085925.
 XX 18-MAY-1998; 98US-0085927.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
 PI Lafleur DM, Endress GA, Ebner R;
 XX WPI; 2000-062296/05.
 XX
 XX New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders -
 XX
 XX Disclosure; Page 440-441; 475pp; English.
 XX
 XX AA265250 to AA265350 represent 97 isolated human secreted protein genes.
 CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
 CC genes. This sequence represents a fragment of one of the human secreted
 CC proteins. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample

FT	XX	/note= "Mature human protein with hydrophobic domain"	Matches 389; Conservative 0; Mismatches 1; Incols U; gaps
FW	XX	W0200149728-A2.	
XX	XX		
PD	XX	12-JUL-2001.	
XX	XX		
PF	XX	28-DEC-2000; 2000WO-JP09359.	
XX	XX		
PR	XX	06-JAN-2000; 2000JP-0000585.	
PR	XX	06-JAN-2000; 2000JP-0000588.	
PR	XX	11-JAN-2000; 2000JP-0002299.	
PR	XX	03-FEB-2000; 2000JP-0026862.	
PR	XX	03-MAR-2000; 2000JP-0058367.	
XX	XX		
PA	XX	(PROT-) PROTEGENE INC.	
PA	XX	(SAGA) SAGAMI CHEM RES CENT.	
PI	XX	Kato S, Kimura T;	
XX	XX		
DR	XX	WPI; 2001-418355/44.	
DR	XX	N-PSDB; AAD12605.	
XX	XX		
PT	XX	Human proteins with hydrophobic domains and the nucleic acids encoding	
PT	XX	them, useful for preventing diagnosing and treating e.g. cancer.	
PT	XX	Alzheimer's and inflammation	
XX	XX		
PS	XX	Claim 1; Page 448-450; 563pp; English.	
XX	XX		
CC	XX	The present sequence is human protein with hydrophobic domain,	
CC	XX	HP10801. The polynucleotide and polypeptide of the invention	
CC	XX	may be used in the prevention, diagnosis and treatment of diseases	
CC	XX	associated with inappropriate polypeptide expression. The	
CC	XX	polynucleotides	
CC	XX	may be used to produce the polypeptide, by inserting the nucleic acids	
CC	XX	into a host cell and culturing the cell to express the protein. The	
CC	XX	polynucleotides and its complementary sequences may also be used as DNA	
CC	XX	probes in diagnostic assays and also used in gene therapy. The	
CC	XX	polypeptides may also be used as antigens in the production of	
CC	XX	antibodies	
CC	XX	and in assays to identify modulators of polypeptide expression and	
CC	XX	activity. The polypeptides and nucleic acids may be used as nutritional	
CC	XX	supplements, to modulate cytokine and cell proliferation activity, to	
CC	XX	modulate immune stimulation or suppression (e.g. for the treatment of	
CC	XX	microbial infections and autoimmune disorders such as multiple	
CC	XX	sclerosis,	
CC	XX	rheumatoid arthritis and insulin-dependent diabetes), to modulate	
CC	XX	haematopoiesis, to modulate tissue growth activity (e.g. for the	
CC	XX	treatment of Parkinson's disease, Huntington's disease and Alzheimer's	
CC	XX	disease), to modulate activin and inhibin activity (e.g. for controlling	
CC	XX	fertility), to modulate chemotactic and chemokinetic activity, to	
CC	XX	modulate haemostatic and thrombolytic activity, to modulate receptor	
XX	XX	ligand activity, to modulate inflammation and to inhibit tumour growth.	
SQ	XX	Sequence 390 AA;	
		Query Match 99.6%; Score 2004; DB 22; Length 390;	
		Best Local Similarity 99.7%; Pred. No. 2.3e-142;	

QY	1	MISLPGFLVTLNRLRFLGLSALAPPSSRAQLQHLPANRLQAVEGEVVLPAYTTLHGEV	60
DB	1	mislpfpvlvnlrrlrlfllglsalappraqqlhlparrlqavgevvlpaytllhgev	60
QY	61	SSSQPWEVPFVMMFFKQKEKEDQVLSYINGVYTSKPGSVLYVSNPRLSLRLEGLQEKD	120
DB	61	sssqpwevpfvmvffkqkqkedqvlsyngvttakpgsalvymparlnlrlrlgllqekd	120
QY	121	SPFYSCSVNQKQKSGHSIKTLELNLVLPAPPSRCLOQVPHVGNVTLSCQSPRSK	180
DB	121	sgpyscavnvqdkgkarghaiktlielnvlpappscrllqgvphvganvclscqgsprsk	180
QY	181	PAVOYQMDROLPSFOTFPFALPDVIRGSLSTNLSSMAGVYVCKARHEVGTACNVTLLE	240
DB	181	pavoyqwdrlpfdqfqtffapaldvirgslstnlssmagvyvckahnevtacnvtlle	240
QY	241	VSTGPGAIVAGAVVGTIVGLIGLILAGLVLLYHRGKALEEPANDIKEDAIAPRTLPPWPKS	300
DB	241	vatgpgaaavagavvgvtivglilaglvlllyhrgkaleepandikedaisprtlpwpsk	300
QY	301	SDTISKNGTSSVTSARALRPPHGGPRRGAUTPTPSLSSQALSPSLPFTTGAAHPQPISP	360
DB	301	sdtsikngtlesvtsaaralrpphggprpgaltptpsalsqalpsprlptdghbpgpiisp	360
QY	361	IPGGVSSGSLSRMGAVFVMPVPAQSQAGSIV	390
DB	361	ipggvasgslsrmgavfvmvmpadsgagelv	390
RESULT 13			
AAB90818			
ID	AAB90818	standard; Protein; 390 AA.	
XX			
AC	AAB90818;		
XX			
DT	15-JUN-2001	(first entry)	
XX			
DE	Human shear stress-response protein SEQ ID NO: 144.		
XX			
KW	Human; shear stress-response protein, vascular disease;		
KW	arteriosclerosis.		
OS	Homo sapiens.		
XX			
FN	W0200125427-A1.		
XX			
PD	12-APR-2001.		
XX			
PF	02-OCT-2000; 2000WO-JP06840.		
XX			
PR	01-OCT-1999; 99JP-0280976.		
XX			
PA	{KYOW } KYOWA HAKKO KOGYO KK.		
PA	{NOJII/} NOJIMA H.		
XX			

[illegible]

QY 241 VSTPGAAVVAAGVGTIVGLGLAGLVLYHRRSKALEBPANDIKEDALAPRTLEWPKS 300
 |||||
 DB 241 veltpgaavvaeevvgtlvgllaglvlyhrrgkaleepandikedalprtllwpks 300
 |||||
 QY 301 SPTISKRGITUSVTSARALRPHGPPRPQALTPPSLSSQALLPSR-----LPTTQGA 353
 |||||
 DB 301 edtlkxgtlsvtaaxalrtpbkpprpgaltpcplesqalpsrphahdiwgpelnlp 360
 |||||
 QY 354 HP 355
 ||
 DB 361 hp 362

Search completed: August 19, 2002, 17:09:06
 Job time: 3349 sec

OM protein - protein search, using sw model

Run on: August 19, 2002, 16:16:12 ; Search time 24.02 Seconds
(without alignments)

396.585 Million cell

updates/sec

Title: US-09-902-759-39

Perfect score:

Sequence: 1 MISLPGPLVTNLLRFLFLGL.....SRMGAVPVMVPAQSQAGSLV 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 04

Maximum Match 100

Listing first 45

Database : Issued_Patents_A

1: /cgn2_6/ptoda

2: /cgn2_6/ptoda

3: /cgn2_6/ptoda

4: /cgn2_6/ptoda

5: /cgn2_6/ptoda

6: /cgn2_6/ptoda

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2012	100.0	390	2	US-08-979-424-1
2	353.5	17.6	365	4	US-08-928-383B-26
3	346	17.2	365	4	US-08-928-383B-26
4	345.5	17.2	365	4	US-08-928-383B-23
5	343	17.0	365	2	US-08-979-424-3
6	343	17.0	365	4	US-08-979-424-3
7	297	14.8	319	1	US-08-957-495B-22
8	297	14.8	319	4	US-08-957-495B-22
9	290.5	14.4	365	4	US-08-968-051A-22
10	289.5	14.4	387	4	US-08-928-383B-24
11	289.5	14.4	387	4	US-08-928-383B-24
12	258	12.8	318	4	US-08-975-928-2
13	258	12.8	318	4	US-08-968-051A-32
14	258	12.8	318	4	US-08-968-051A-32
15	258	12.8	318	4	US-08-968-051A-32
16	258	12.8	318	4	US-08-968-051A-32
17	258	12.8	318	4	US-08-968-051A-32
18	258	12.8	318	4	US-08-968-051A-32
19	258	12.8	318	4	US-08-968-051A-32
20	258	12.8	318	4	US-08-968-051A-32
21	258	12.8	318	4	US-08-968-051A-32
22	258	12.8	318	4	US-08-968-051A-32
23	258	12.8	318	4	US-08-968-051A-32
24	258	12.8	318	4	US-08-968-051A-32
25	258	12.8	318	4	US-08-968-051A-32
26	258	12.8	318	4	US-08-968-051A-32
27	258	12.8	318	4	US-08-968-051A-32
28	258	12.8	318	4	US-08-968-051A-32
29	258	12.8	318	4	US-08-968-051A-32
30	258	12.8	318	4	US-08-968-051A-32
31	258	12.8	318	4	US-08-968-051A-32
32	258	12.8	318	4	US-08-968-051A-32
33	258	12.8	318	4	US-08-968-051A-32
34	258	12.8	318	4	US-08-968-051A-32
35	258	12.8	318	4	US-08-968-051A-32
36	258	12.8	318	4	US-08-968-051A-32
37	258	12.8	318	4	US-08-968-051A-32
38	258	12.8	318	4	US-08-968-051A-32
39	258	12.8	318	4	US-08-968-051A-32
40	258	12.8	318	4	US-08-968-051A-32
41	258	12.8	318	4	US-08-968-051A-32
42	258	12.8	318	4	US-08-968-051A-32
43	258	12.8	318	4	US-08-968-051A-32
44	258	12.8	318	4	US-08-968-051A-32
45	258	12.8	318	4	US-08-968-051A-32
46	258	12.8	318	4	US-08-968-051A-32
47	258	12.8	318	4	US-08-968-051A-32
48	258	12.8	318	4	US-08-968-051A-32
49	258	12.8	318	4	US-08-968-051A-32
50	258	12.8	318	4	US-08-968-051A-32
51	258	12.8	318	4	US-08-968-051A-32
52	258	12.8	318	4	US-08-968-051A-32
53	258	12.8	318	4	US-08-968-051A-32
54	258	12.8	318	4	US-08-968-051A-32
55	258	12.8	318	4	US-08-968-051A-32
56	258	12.8	318	4	US-08-968-051A-32
57	258	12.8	318	4	US-08-968-051A-32
58	258	12.8	318	4	US-08-968-051A-32
59	258	12.8	318	4	US-08-968-051A-32
60	258	12.8	318	4	US-08-968-051A-32
61	258	12.8	318	4	US-08-968-051A-32
62	258	12.8	318	4	US-08-968-051A-32
63	2				

17

12	176	8.7	299	4	US-09-180-350-331	Sequence 224, Appl
13	176	8.7	299	4	US-09-462-270-2	Sequence 2, Appl
14	171	8.5	299	4	US-09-188-930-189	Sequence 189, Appl
15	158.5	7.9	501	2	US-08-040-095-31	Sequence 31, Appl
16	153.5	7.6	344	2	US-08-602-725-34	Sequence 34, Appl
17	152	7.6	1101	3	US-08-986-485-2	Sequence 2, Appl
18	134	6.7	607	2	US-08-752-307B-12	Sequence 12, Appl
19	133	6.6	321	6	5169835-17	Patent No. 5169835
20	133	6.6	464	2	US-08-603-725-32	Sequence 32, Appl
21	133	6.6	642	1	US-08-414-657D-1	Sequence 1, Appl
22	133	6.6	698	2	US-08-602-725-36	Sequence 36, Appl
23	133	6.6	734	2	US-08-389-459A-17	Sequence 17, Appl
24	133	6.6	734	3	US-08-987-867A-17	Sequence 17, Appl
25	132.5	6.6	252	2	US-08-414-657D-58	Sequence 58, Appl
26	132.5	6.6	287	2	US-08-414-657D-46	Sequence 46, Appl
27	132.5	6.6	304	2	US-08-414-657D-45	Sequence 44, Appl
28	132.5	6.6	308	2	US-08-414-657D-44	Sequence 46, Appl
29	132.5	6.6	325	2	US-08-414-657D-2	Sequence 2, Appl
30	132.5	6.6	325	2	US-08-414-657D-41	Sequence 41, Appl
31	132.5	6.6	338	2	US-08-414-657D-60	Sequence 60, Appl
32	132.5	6.6	1241	4	US-09-040-774-2	Sequence 2, Appl
33	132	6.6	828	1	US-08-261-304-2	Sequence 2, Appl
34	130	6.5	1091	3	US-08-986-485-5	Sequence 5, Appl
35	130	6.4	917	1	US-08-245-235-2	Sequence 2, Appl
36	129	6.4	917	1	US-08-481-130-2	Sequence 2, Appl
37	129	6.4	917	1	US-08-656-984A-2	Sequence 2, Appl
38	129	6.4	917	1	US-08-485-604-2	Sequence 2, Appl
39	129	6.4	917	2	US-08-487-595-2	Sequence 2, Appl
40	128.5	6.4	252	2	US-08-414-657D-57	Sequence 57, Appl
41	128.5	6.4	287	2	US-08-414-657D-49	Sequence 49, Appl
42	128.5	6.4	310	2	US-08-414-657D-45	Sequence 45, Appl
43	128.5	6.4	315	2	US-08-414-657D-47	Sequence 47, Appl
44	128.5	6.4	338	2	US-08-414-657D-42	Sequence 42, Appl
45	128.5	6.4	338	2	US-08-414-657D-43	Sequence 43, Appl

ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```


RESULT 2

US-08-928-383B-26
; Sequence 26, Application US/08928383B
; Patent No. 6210921
; GENERAL INFORMATION:
; APPLICANT: Robert W. Pinberg, Jeffrey M. Bergelson,
; APPLICANT: and Marshall S. Horwitz
; TITLE OF INVENTION: CAR, A NO. 6210921el Cxsackievirus and Adenovirus
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,383B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,100
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-383B-26

US-08-928-383B-26
; Sequence 26, Application US/08928383B
; Patent No. 6210921
; GENERAL INFORMATION:
; APPLICANT: Robert W. Pinberg, Jeffrey M. Bergelson,
; APPLICANT: and Marshall S. Horwitz
; TITLE OF INVENTION: CAR, A NO. 6210921el Cxsackievirus and Adenovirus
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,383B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,100
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-383B-26

Query Match 100.0%; Score 2012; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.7e-168;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MISLPGPLVTNLLRFLFLGLSALAPPSPRAQLQLHPANRLOAVEGGEVLPANWTLHGEV 60
DB 1 MISLPGPLVTNLLRFLFLGLSALAPPSPRAQLQLHPANRLOAVEGGEVLPANWTLHGEV 60
QY 61 SSSQFNEVPVWFFPKQKEDQVLSYINGVTTSPKGVSLVSMSPRNLRLSLEGLQEKD 120
DB 61 SSSQFNEVPVWFFPKQKEDQVLSYINGVTTSPKGVSLVSMSPRNLRLSLEGLQEKD 120
QY 121 SGPYSCSVAVQDKQKSRGHSIKTLELVLPVPPAPPSCRQLQGVPHVGNVTLSCQSPRSK 180
DB 121 SGPYSCSVAVQDKQKSRGHSIKTLELVLPVPPAPPSCRQLQGVPHVGNVTLSCQSPRSK 180
QY 181 PAVQYQDRQLPSQFTFPALDVLVIRGSLNLTSSNAGVYVCAHNEVGTACQNVTL 240
DB 181 PAVQYQDRQLPSQFTFPALDVLVIRGSLNLTSSNAGVYVCAHNEVGTACQNVTL 240
QY 241 VSTGPGAAVAGVVGTVLGLGLAGLVLYHRRKGALEPANDIKEDATAPETLPWPKS 300
DB 241 VSTGPGAAVAGVVGTVLGLGLAGLVLYHRRKGALEPANDIKEDATAPETLPWPKS 300
QY 301 SDTISKNGTLLSVTSARALRPHGPPRFGALFTPSLSSQALPSPRLFTDGAHPQIPSP 360
DB 301 SDTISKNGTLLSVTSARALRPHGPPRFGALFTPSLSSQALPSPRLFTDGAHPQIPSP 360

Query Match 17.6%; Score 353.5; DB 4; Length 365;
Best Local Similarity 27.8%; Pred. No. 4.7e-23;
Matches 113; Conservative 71; Mismatches 156; Indels 67; Gaps 15;
QY 9 VTNLLRFLFL-GLSALAPPSPRAQLQLHPANRLOAVEGGEVLPANWTLHGEVSSSQPWE 67
DB 1 MARLLCFLVLCGLADFT----SGLSITTPQRIEKAKGATAYLPCKFTLSPE--DQGPLD 54

Search completed: August 19, 2002, 17:09:51
Job time: 3219 sec

OM protein - protein search, using sw model

Run on: August 19, 2002, 16:21:32 ; Search time 42.75 Seconds
(without alignments)
876,604 Million cell

updates/sec

Title: US-09-902-759-39
Perfect score: 2012
Sequence: 1 MISLPGPLVTLRLPLPLGL.....SIMGAVPVMVPAQSQAGSLV 390

Scoring table: BLOS62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	176	8.7	299	2 SS6749	junctional adhesio
2	162.5	8.1	344	2 A27681	non-specific cross-
3	158.5	7.9	847	2 JH0371	B-cell adhesion pr
4	155.5	7.7	3707	2 S18252	heparan sulfate pr
5	154.5	7.7	4391	2 A38096	perlecan precursor
6	153.5	7.6	647	2 A35648	B-cell adhesion pr
7	153.5	7.6	1040	2 A49356	transient axonal
8	152.5	7.6	521	2 S34338	biliary glycoprote
9	152	7.6	483	2 T17346	hypothetical prote
10	151.5	7.5	1036	2 S22383	axinin 1 precursor
11	151.5	7.5	4162	2 T42633	connectin/titin -
12	150.5	7.5	521	2 JCI508	biliary glycoprote

14	150	7.5	868	2	A46512	U244 homolog/a ym	
15	149.5	7.4	341	2	JCI512	biliary glycoprote	
16	147	7.3	5175	2	T20992	hypothetical prote	
17	147	7.3	5198	2	T43290	hemictin precurs	
18	144	7.2	349	2	A34815	carcinoembryonic	
a	19	143	7.1	862	2	I49583	differentiation an
	20	142	7.1	278	2	A39037	carcinoembryonic
a	21	141.5	7.0	26926	1	I38344	titin, cardiac mus
	22	141	7.0	458	2	JCI509	biliary glycoprote
	23	139	6.9	458	1	WMSR1	biliary glycoprote
	24	138.5	6.9	495	2	A55181	pregnancy-specific
	25	138.5	6.9	1323	2	PNO568	connectin 3B - chi
	26	138	6.9	278	2	JCI506	biliary glycoprote
	27	138	6.9	419	2	B54312	pregnancy-specific
	28	137	6.8	7962	2	I38346	elastic titin - hu
	29	136.5	6.8	518	2	JC4024	poliovirus recepto
	30	135.5	6.7	426	2	C55181	pregnancy-specific
	31	135.5	6.7	426	2	B35334	pregnancy-specific
	32	135	6.7	428	2	JS0032	pregnancy-specific
	33	135	6.7	428	2	I57486	pregnancy-specific
	34	134	6.7	240	2	JC4121	pregnancy-specific
	35	134	6.7	436	2	B55181	pregnancy-specific
	36	133	6.6	321	2	JH0395	biliary glycoprote
	37	133	6.6	351	2	JH0396	biliary glycoprote
	38	133	6.6	417	2	JH0394	biliary glycoprote
	39	133	6.6	419	2	A36109	pregnancy-specific
	40	133	6.6	464	2	C30127	transmembrane carc
	41	133	6.6	526	1	A32164	biliary glycoprote
	42	133	6.6	702	2	A36319	carcinoembryonic
a	43	132.5	6.6	166	2	A33402	pregnancy-specific
	44	132.5	6.6	338	2	JC4776	limbic-system-asso
	45	132.5	6.6	338	2	JCI238	oploid-binding pro

ALIGNMENTS

RESULT 1	SS6749	junctional adhesion molecule precursor - human
	N/Alternate names:	F11 platelet antigen; platelet adhesion molecule PAM-1;
	platelet F11 receptor	
	C/Species:	Homo sapiens (man)
	C/Date:	27-Oct-1995 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
	R/Ozaki, H.; Ishii, K.; Horiuchi, A.; Arai, H.; Kawamoto, T.; Okawa, K.;	
	Iwanatsu, A.; Kita, T.	
	J. Immunol.	163, 553-557, 1999
	A/Title:	Cutting edge: combined treatment of TNF-alpha and IFN-gamma causes
	A/Reference number:	A59406; MUID:99323940; PMID:10395639
	A/Accession:	A59406
	A/Status:	preliminary
	A/Molecule type:	DNA

A:Residues: 1-299 <OZA>
 A:Cross-references: GB:AAD42050; NID:95326797; PIDN:AAD42050.1
 R:Naik, U.P.; Ehrlich, Y.H.; Kornecki, E.
 Biochem. J. 310, 155-162, 1995
 A:Title: Mechanisms of platelet activation by a stimulatory antibody:
 cross-linking of a novel platelet receptor for monoclonal antibody F11 with
 the Fc-gamma-RII receptor.
 A:Reference number: S56749; PMID:95374438; PMID:7646439
 A:Accession: S56749
 A:Molecule type: protein
 A: ; R e s i d u e s :
 28-49, 'X', 51-53; 62-73, 'E', 75-103; 123, 'F', 125-130; 'FDRDXTIYINXY', 'LT', 206, 'X'
 '208, 'Q', <NAI>
 A>Note: the order of the peptides other than the amino terminus was not
 determined
 C:Genetics:
 A:Gene: JAM
 C:Keywords: glycoprotein; phosphoprotein; platelet aggregation; platelet
 membrane
 F11-25/Domain: signal sequence #status predicted <SIG>
 F126-299/Product: junctional adhesion molecule #status predicted <MAT>

Query Match 8.7%; Score 176; DB 2; Length 299;
 Best Local Similarity 25.7%; Pred. No. 0.00014;
 Matches 71; Conservative 46; Mismatches 129; Indels 30; Gaps 12;

QY	9	VTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSSQPWEV	68
Db	7	VERKLLCLFTAILLCSLALGSGVTSHSSEPEVRIPENRPFVLSAYS	57
QY	69	PFVMFFKQKEKDQVLSYINGVTTSKPGVSLVSWPSRNLRLGLEKQKSGPYSCSV	128
Db	58	PRVENKPDQD-TTRLVCYNKITASYE--DRVTFLEPT--GITPKSVTRDTGTTCMV	111
QY	129	NVQDKQKSGRHSIKTLEINLVFPAPPSCNLOGVPHVGANVTLSQSPRSKPAVOYQWD	188
Db	112	S--EEGNGSYG-EVK-VKLIIVLPSPKPTNIPSSATIGNRAVLTCSEQDGSPPSEYTFW	167
QY	189	RQ---LP-----SFQTF--PAPALDVIRGSLSLTLSSMAGVYVCKAHNEVGTQA-CNVT	238
Db	168	KGIQVMPNPKSTRAPNSVSNFTTGLVDFPLASDTCGEVSCAENGVTPTMSNAV	227
QY	239	LEVSTGPGAAVAGAVGTLVGLGLLA-GIVLLYHR	273
Db	228	RMEAVERNVGVRAVLTLILGILVFGINPAYSR	263

Search completed: August 19, 2002, 17:10:57
 Job time: 2965 sec

OM protein - protein search, using sw model

Run on: August 19, 2002, 17:11:02 ; Search time 24.06 Seconds
(without alignments)
627.624 Million cell

updates/sec

Title: US-09-902-759-39

Perfect score: 2012

Sequence: 1 MISLPGPLVTLRLFLGL.....SEMGVPMVPAQAGSIV 390

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:

105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters:

105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33.5	17.6	365	1	P97792 mus musculus
2	34.3	17.0	365	1	FX3310 homo sapien
3	297	14.8	319	1	A33_HUMAN
4	212	10.5	298	1	JAM2_HUMAN
5	176	8.7	299	1	JAM1_HUMAN
6	168.5	8.4	300	1	JAM1_MOUSE
7	167.5	8.3	298	1	JAM1_BOVIN
8	163	8.1	344	1	CEA5_HUMAN
9	158.5	7.9	4393	1	PCBM_MOUSE
10	155.5	7.7	3707	1	PCBM_MOUSE
11	153.5	7.6	847	1	CD22_HUMAN
12	153.5	7.6	1040	1	CD22_HUMAN
13	151.5	7.5	1036	1	AXOI_CHICK
14	150.5	7.5	521	1	CEA1_MOUSE
15	150.5	7.5	1040	1	AXOI_MOUSE
16	146	7.3	515	1	PVR1_PIG
17	145	7.2	517	1	PVR1_HUMAN

19	143	7.1	862	1	Q222_MOUSE
20	142.5	7.1	1709	1	SN_HUMAN
21	138.5	6.9	515	1	PVR1_MOUSE
22	138	6.9	348	1	KILO_RAT
23	138	6.9	419	1	PSGA_HUMAN
24	135.5	6.7	426	1	PSGB_HUMAN
25	135	6.7	428	1	PSG3_HUMAN
26	134.5	6.7	337	1	GS5A_CHICK
27	133	6.6	526	1	CEA1_HUMAN
28	133	6.6	702	1	CEA5_HUMAN
29	132.5	6.6	338	1	LAMP_HUMAN
30	132.5	6.6	345	1	OPCM_HUMAN
31	132.5	6.6	345	1	OPCM_RAT
32	132	6.6	252	1	CEA3_HUMAN
33	131.5	6.5	349	1	LACH_SCEAM
34	131.5	6.5	538	1	PVR2_HUMAN
35	131	6.5	519	1	ECTO_RAT
36	130.5	6.5	837	1	NCM2_MOUSE
37	130	6.5	761	1	ICM2_HUMAN
38	130	6.5	917	1	ICM5_MOUSE
39	129.5	6.4	345	1	OPCM_BOVIN
40	128.5	6.4	338	1	LAMP_RAT
41	128.5	6.4	837	1	NCM2_HUMAN
42	128	6.4	338	1	LAMP_CHICK
43	128	6.4	530	1	PVR2_MOUSE
44	126.5	6.3	404	1	RAGE_HUMAN
45	126	6.3	335	1	PSG5_HUMAN

ALIGNMENTS

RESULT 1

CXAR_MOUSE

ID CXAR_MOUSE STANDARD; PRT; 365 AA.

AC P97792; 009052;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, last sequence update)

DT 01-MAR-2002 (Rel. 41, last annotation update)

DE Coxsackievirus and adenovirus receptor homolog precursor (mCAR).

GN CXADR OR CAR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Liver;

RX MEDLINE=9710109; PubMed=9036860;

RA Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,

RA Krithivas A., Hong J.S., Horwitz M.S., Crowell R.L., Finberg R.W.;

RT "Isolation of a common receptor for Coxsackie B viruses and

RL Science 275:1320-1323(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C3H/MAI;

RX MEDLINE=97250541; PubMed=9096397;
 RA Tomko R.P., Xu R., Philipson L.;
 RT "HCR and MCR: the human and mouse cellular receptors for subgroup C
 RT adenoviruses and group B coxsackieviruses";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RA Bergelson J.M., Krithivas A., Crowell T.L., Finberg R.W.;
 RT "The murine CAR homologue (MCR) is a receptor for coxsackie B
 RT viruses and adenoviruses";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 C
 CC This SWISS-PROT entry is copyright. It is produced through a
 CC collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no
 CC modified and this statement is not removed. Usage by and for
 CC commercial
 CC entities requires a license agreement (See
 CC <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 C
 DR EMBL; Y10320; CAA71368.1; -;
 DR EMBL; U90715; AAC53148.1; -;
 DR EMBL; Y11929; CAA72679.1; -;
 DR MGD; MGI:1201679; Cxadr.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR003600; Ig_Like.
 DR Pfam; PF000047; Ig_2.
 DR SMART; SM00410; Ig_Like; 1.
 DR SMART; SM00408; IGC2; 1.
 KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;
 KW Repeat.
 FT SIGNAL
 FT CHAIN 1 19 POTENTIAL.
 FT CHAIN 20 365 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR
 FT CHAIN 20 365 HOMOLOG.
 FT DOMAIN 20 237 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 238 258 POTENTIAL.
 FT DOMAIN 259 365 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 34 127 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 155 219 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 41 120 BY SIMILARITY.
 FT DISULFID 162 212 BY SIMILARITY.
 FT CARBOHYD 106 106 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 340 365 VAAPELSEMGAVPMIPAQSGDSIV -> FKAYKTDGIT
 FT (IN REF. 2 AND 3).

SQ SEQUENCE 365 AA; 39947 MW; 54455945ZAJ46ZAZ UCUGA;
 Query Match 17.6%; Score 353.5; DB 1; Length 365;
 Best Local Similarity 27.8%; Pred. No. 1.8e-16;
 Matches 113; Conservative 71; Mismatches 156; Indels 67; Gaps
 15;
 QY 9 VTNLRLFLPL-GLSALAPPSPRAQLQHLHPANRLQAVEGGEVILPWWYTLHGRVSSQPWE 67
 DB 1 MARLLCFVLGGIADFT----SGLSITTPRRIEAKGKGTAYLPCKFTLSPE--DQGPLD 54
 QY 68 VPFVNVFFKQKE--DQVLSYINGVTVSKPGVSLVY-----SMSPRNL 109
 DB 55 IE---WLISPSDNQIVDQVILYSG-----DKIYDNYVPDLKGRVHFTSNDVKSQDA 103
 QY 110 SURLEGLQEKDSQFVSCSVNVQDKKGRGHSIKTLELVLDVAPPPCRLQGVPHVGAN 169
 DB 104 SINVTNLQSLDITGTYQCKVK-----KAPGVANKKFLTLVLVXPSTGTCFVDGSEIAMD 157
 QY 170 VTLSQSPRSKPAVOYQWDRQLSPFTFPAPAL-DVIRGSLTNLSLSSMAGVTVCKAIN 228
 DB 158 FKUKCEPKESGLPQFEW-QKUSDSQTPPTWLAEMTSFVIVSNASSEYGTCTVQVN 216
 QY 229 EVGTACQNVTLIE-VSTGFGAAVAVAGVGTUIGLGLAGLVLYHER---GKALEEPAND 284
 DB 217 RVGSDQCHLRLDVVPPSNRAGTTAGAVIGTLLALVIGALLFCCHRRKREKEVEKVED 276
 QY 285 IKEDAIAPRTLPHWPKSDTSIKNGTSSVTSARALPPHGPFPREGALTPTFSLSSQALPS 344
 DB 277 IRED-----VPPPKSRTSTARSYIGSNHSSL-----GWSMSPSNMEGYSKTYQ 318
 QY 345 PRLLPTTDGAH-POPISPIPGVSSGSLRMGAVPMVPAQSOAGSLV 390
 DB 319 NQVPSDFERAQSPPTLAPAKVAAPNLSRMGAVPMVPAQSKOGSIV 365

Search completed: August 19, 2002, 17:18:28
 Job time: 446 sec

OM protein - protein search, using sw model

Run on: August 19, 2002, 17:09:12 ; Search time 67.26 Seconds
(without alignments)
updates/sec 1003.093 Million cell

Title: US-09-902-759-39
Perfect score: 2012
Sequence: 1 MISLPGVTLNLLRFLFLGL.....SRMGAVFVMPAQSGSLV 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL19:*
 - 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_muc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_virtebrate.*
 - 14: sp_unclassified.*
 - 15: sp_virus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description

2	2009	99.9	390	4	Q96TSU	Q96TSU homo sapien
3	1931	96.0	394	6	Q9SK13	Q9SK13 macaca fasc
4	1397	69.4	330	11	Q925F2	Q925F2 mus musculus
5	641.5	31.9	204	11	Q9DJ12	Q9DJ12 mus musculus
6	343	17.0	366	11	Q9DBJ8	Q9DBJ8 mus musculus
7	341.5	17.0	372	13	Q90Y50	Q90Y50 brachydanio
8	330	16.4	358	11	Q9R065	Q9R065 rattus norv
9	324.5	16.1	337	4	Q961Q7	Q961Q7 homo sapien
10	320.5	15.9	325	4	Q95791	Q95791 homo sapien
11	316	15.7	407	11	Q9D2J4	Q9D2J4 mus musculus
12	307.5	15.3	373	4	Q9HGB4	Q9HGB4 homo sapien
13	290	14.4	284	4	Q9NK42	Q9NK42 homo sapien
14	287	14.3	373	11	Q920S5	Q920S5 mus musculus
15	279.5	13.9	318	13	Q91664	Q91664 xenopus lae
16	277.5	13.8	300	11	Q9D9J0	Q9D9J0 mus musculus
17	276.5	13.7	300	11	Q9DA22	Q9DA22 mus musculus
18	272	13.5	352	11	Q91W66	Q91W66 mus musculus
19	269	13.4	304	11	Q9CVA4	Q9CVA4 mus musculus
20	267.5	13.3	328	11	Q92109	Q92109 mus musculus
21	267.5	13.3	344	11	Q9R067	Q9R067 rattus norv
22	266	13.2	335	13	Q9PWR4	Q9PWR4 gallus gall
23	264.5	13.1	344	4	Q9UKV4	Q9UKV4 homo sapien
24	263	13.1	335	13	Q9YGH1	Q9YGH1 gallus gall
25	262.5	13.0	259	4	Q95532	Q95532 homo sapien
26	259	12.9	319	11	Q922D5	Q922D5 mus musculus
27	258	12.8	319	6	Q9TU80	Q9TU80 canis famil
28	258	12.8	319	11	Q9JKAS	Q9JKAS mus musculus
29	257	12.8	248	11	Q9D0T4	Q9D0T4 mus musculus
30	256	12.7	319	6	Q9TU79	Q9TU79 sus scrofa
31	253	12.6	335	13	Q9YGV5	Q9YGV5 gallus gall
32	221	11.0	181	13	Q91665	Q91665 xenopus lae
33	196	9.7	577	11	Q9D221	Q9D221 mus musculus
34	186.5	9.3	977	4	Q96RD9	Q96RD9 homo sapien
35	183	9.1	300	11	Q9JHY1	Q9JHY1 rattus norv
36	176	8.7	298	11	Q9J159	Q9J159 mus musculus
37	171	8.5	381	4	Q9Y4A4	Q9Y4A4 homo sapien
38	169.5	8.4	309	4	Q96FL1	Q96FL1 homo sapien
39	169.5	8.4	310	4	Q9EX67	Q9EX67 homo sapien
40	168.5	8.4	430	4	Q15600	Q15600 mus sapien
41	166.5	8.3	310	11	Q9DBB7	Q9DBB7 mus musculus
42	166	8.3	310	11	Q9EPK4	Q9EPK4 mus musculus
43	166	8.3	310	11	Q9DLM9	Q9DLM9 mus musculus
44	164	8.2	538	4	Q9NWQ7	Q9NWQ7 homo sapien
45	163.5	8.1	512	4	Q96DN8	Q96DN8 homo sapien

ALIGNMENTS

RESULT	1
Q96AP7	
ID	Q96AP7 PRELIMINARY;
AC	Q96AP7; PRT; 390 AA.
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	HYPOTHETICAL 41.2 KDA PROTEIN.

CS. Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2123298; PubMed=11279107;
 RA Hirata K.-I., Ishida T., Penta K., Rezaee M., Yang E., Wohlgemuth J.,
 RA Quatermous T.;
 RT "Cloning of an immunoglobulin family adhesion molecule selectively
 RT expressed by endothelial cells.";
 RL J. Biol. Chem. 276:16223-16231(2001).
 DR EMBL; AF361746; AAKS1065.1; --
 SQ SEQUENCE 390 AA; 41176 MW; C5E3F302F41B6E8C CRC64;

Query Match 100.0%; Score 2012; DB 4; Length 390;
 Best Local Similarity 100.0%; Pred. No. 9.6e-156; Mismatches 0; Indels 0; Gaps 0;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MISLPGVLTNLRFLFLGLSALAPPSPRAQLQLHPANRLOAVGGVVLPAWYTLHGEV 60
 DB 1 MISLPGVLTNLRFLFLGLSALAPPSPRAQLQLHPANRLOAVGGVVLPAWYTLHGEV 60
 QY 61 SSSQPEVFFVWFFFKQEKEDQVLSYNGVTTSKPGVSLVYSPNSLSLEGLQEKD 120
 DB 61 SSSQPEVFFVWFFFKQEKEDQVLSYNGVTTSKPGVSLVYSPNSLSLEGLQEKD 120
 QY 121 SGPYSCVNVQDKGSRGHSIKTLELVLPAPPSCRLQGVPHGVANVTLSQSPRSK 180
 DB 121 SGPYSCVNVQDKGSRGHSIKTLELVLPAPPSCRLQGVPHGVANVTLSQSPRSK 180
 QY 181 PAVQYQWDRQLPSPQTFPPAPALDIVRGSLSLTLNLSMAGVYVCKAHNEVGTACQNTLE 240
 DB 181 PAVQYQWDRQLPSPQTFPPAPALDIVRGSLSLTLNLSMAGVYVCKAHNEVGTACQNTLE 240
 QY 241 VSTPGGAAYVAGAVGTGLVGLLGLAGLVLLYHRRGKALSEPANDIKEDALAPRTLWPKS 300
 DB 241 VSTPGGAAYVAGAVGTGLVGLLGLAGLVLLYHRRGKALSEPANDIKEDALAPRTLWPKS 300
 QY 301 SDTISKNGTLSSVTSARALRPHGPGRPCALPTPTSLSSOALPSPRLPTDGNHPQIPSP 360
 DB 301 SDTISKNGTLSSVTSARALRPHGPGRPCALPTPTSLSSOALPSPRLPTDGNHPQIPSP 360
 QY 361 IPGVSSSGLSRMGAVFVWVPAQSQAGSLV 390
 DB 361 IPGVSSSGLSRMGAVFVWVPAQSQAGSLV 390

RESULT 2
 Q96T50 PRELIMINARY; PRT; 390 AA.
 AC Q96T50;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ENDOTHelial CELL-SELECTIVE ADHESION MOLECULE.
 GN ESAM.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2123298; PubMed=11279107;
 RA Hirata K.-I., Ishida T., Penta K., Rezaee M., Yang E., Wohlgemuth J.,
 RA Quatermous T.;
 RT "Cloning of an immunoglobulin family adhesion molecule selectively
 RT expressed by endothelial cells.";
 RL J. Biol. Chem. 276:16223-16231(2001).
 DR EMBL; AF361746; AAKS1065.1; --
 SQ SEQUENCE 390 AA; 41208 MW; C5E3EBBSF41B6E8C CRC64;

Query Match 99.9%; Score 2009; DB 4; Length 390;
 Best Local Similarity 99.7%; Pred. No. 1.7e-155; Mismatches 389; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MISLPGVLTNLRFLFLGLSALAPPSPRAQLQLHPANRLOAVGGVVLPAWYTLHGEV 60
 DB 1 MISLPGVLTNLRFLFLGLSALAPPSPRAQLQLHPANRLOAVGGVVLPAWYTLHGEV 60
 QY 61 SSSQPEVFFVWFFFKQEKEDQVLSYNGVTTSKPGVSLVYSPNSLSLEGLQEKD 120
 DB 61 SSSQPEVFFVWFFFKQEKEDQVLSYNGVTTSKPGVSLVYSPNSLSLEGLQEKD 120
 QY 121 SGPYSCVNVQDKGSRGHSIKTLELVLPAPPSCRLQGVPHGVANVTLSQSPRSK 180
 DB 121 SGPYSCVNVQDKGSRGHSIKTLELVLPAPPSCRLQGVPHGVANVTLSQSPRSK 180
 QY 181 PAVQYQWDRQLPSPQTFPPAPALDIVRGSLSLTLNLSMAGVYVCKAHNEVGTACQNTLE 240
 DB 181 PAVQYQWDRQLPSPQTFPPAPALDIVRGSLSLTLNLSMAGVYVCKAHNEVGTACQNTLE 240
 QY 241 VSTPGGAAYVAGAVGTGLVGLLGLAGLVLLYHRRGKALSEPANDIKEDALAPRTLWPKS 300
 DB 241 VSTPGGAAYVAGAVGTGLVGLLGLAGLVLLYHRRGKALSEPANDIKEDALAPRTLWPKS 300
 QY 301 SDTISKNGTLSSVTSARALRPHGPGRPCALPTPTSLSSOALPSPRLPTDGNHPQIPSP 360
 DB 301 SDTISKNGTLSSVTSARALRPHGPGRPCALPTPTSLSSOALPSPRLPTDGNHPQIPSP 360
 QY 361 IPGVSSSGLSRMGAVFVWVPAQSQAGSLV 390
 DB 361 IPGVSSSGLSRMGAVFVWVPAQSQAGSLV 390

RESULT 3
 Q95KI3 PRELIMINARY; PRT; 390 AA.
 AC Q95KI3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 40.9 KDA PROTEIN.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
RM	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=SMISS WEBSTER/NIH;
RX	MEDLINE=21238298; PubMed=11279107;
RA	Hirata K.-I., Ishida T., Penta K., Rezaee M., Yang E., Wohlgenuth J., Quatermous T., "Cloning of an immunoglobulin family adhesion molecule selectively expressed by endothelial cells.," J. Biol. Chem. 276:16223-16231(2001). ENBL: AF361882; AAK51504.1; -. SEQUENCE 394 AA, 41810 MW; 3D2B35493A2227D_CRC64;

Query Match 69.4%; Score 1397; DB 11; Length 394;
Best Local Similarity 72.3%; Pred. No. 1.1e-105;
Matches 285; Conservative 33; Mismatches 72; Indels 4; Gaps 3;

[illegible]

RESULT 5
Q9D712 PRELIMINARY; PRT; 204 AA.
ID Q9D712 AC Q9D712;
DT 01-JUN-2001 (T=EMBLrel. 17. Created)

OC	Cercopithecinae; Macaca.
NCBI_TaxId=9541;	OX
[1]	
SEQUENCE FROM N.A.	RN
TISSUE=TEMPORAL LOBE RIGHT;	RP
RC	RC
Osada N., Hida M., Kuenda J., Tanuma R., Iseki K., Terao K.,	RD
Suzuki Y., Sugano S., Hashimoto K.;	RA
RT Isolation of full-length cDNA clones from macaque brain	RT
libraries.;	RT
Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.	RL
EMBL; AB060855; BAB46874.1; -	DR
hypothetical protein.	KW
Hypothetical protein.	SO
SEQUENCE 390 AA. 40946 MW: CDBF63E2BD464EF5 CRC64;	SO

```

Query Match      96.0%; Score 1931; DB 6; Length 390;
Best Local Similarity 96.4%; Pred. No. 378-149;
Matches 376; Conservative 6; Mismatches 8; Indels 0; Gaps
0;

```

QY	1	MISLPGPLVNTLLRPLFLGLISALAPPRAQLOJHLPAERLQAVGEGEVVLPAMVTLGRV	60
Db	1	MISLPGPLVNTLLRPLFLGLISALAPPRAELQLHLPAHQLOAVGEGEVVLPAMVTLHAEV	60
QY	61	SSSQPWEVPPFMMFFKQKEKEDQVLSYNGVTTSKGSLVYSNPSRNLSLRLEGLQEKD	120
		:	
		:	
Db	61	SSAQGEVPPFMMFFKQKEKEDQVLSYNGVTTSKGSLVYSNPSRNLSLRLEGLQEKD	120
		:	
		:	
QY	121	SGPYSYVNVQDQKQSKRGHSIKITLNLVLPAPPSPCRLOQVPHVGVANVTLSCQSPRSK	180
		:	
Db	121	SGPYSYVNVQDNGQASHGISIKITLNLVLPAPPSPCRLOQVPRVGVANVTLSCQSPRSK	180
		:	
QY	181	PAVOQWDRQLPSFQTPAPPAALDVIKSGLSITLNLSSMAGVYVCKAHNEVGTACQNVILE	240
Db	181	PAVOQWDRQLPSFQTPAPVLDVIRGSLVITLNLSSMAGVYVCKAHNEVGTACQNVILE	240
QY	241	VSTGPGAAVAVAGAVGTGLVGLLQAGVLLVYHRGKALEEPANDIKEDAIAPRTLPMKPS	300
Db	241	VSTGPGAAVAVAGAVGTGLVGLLQAGVLLVYHRGKALEEPANDIKEDAIAPRTLPMKPS	300
QY	301	SDTISKNGTILSSVTSARALRPPHGPFRPGALTTPSLSSQALPSPLPTTGGANPQIPSP	360
Db	301	SDTISKNGTILSSVTSARALRPPHGPFRPGALTTPSLSSQALPSPLPTTGGANPQIPSP	360
QY	361	IPGVSSSGLSRMGAVPVMVPAQSQAGSLV	390
Db	361	IPGVSSSGLSRMGAVPVMVPAQSQAGSLV	390

RESULT	4	
Q925F2		
ID	Q925F2	PRELIMINARY; PRT; 394 AA.
AC	Q925F2;	
DT	01-DEC-2001	(T-EMBLrel. 19, Created)
DT	01-DEC-2001	(T-EMBLrel. 19, Last sequence update)
DT	01-DEC-2001	(T-EMBLrel. 19, Last annotation update)

DT 01-JUN-2001 (TEMPREL. 17, Last sequence update)
 DT 01-DEC-2001 (TEMPREL. 19, Last annotation update)
 DE 23100080SRIK PROTEIN.
 GN EGAM OR 23100080SRIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Akawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayaishizaki Y.
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK009223; BAB26146.1; -.
 DR MGD; MGI:1916774; Esam.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003600; Ig like.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00410; IG like; 1.
 SQ SEQUENCE 204 AA; 22352 MW; 021B29BE2B05F9494 CRC64;

Search completed: August 19, 2002, 17:15:52
 Job time: 400 sec

Query Match 31.9%; Score 641.5; DB 11; Length 204;
 Beat Local Similarity 66.2%; Pred. No. 1.5e-44;
 Matches 129; Conservative 25; Mismatches 38; Indels 3; Gaps 2;
 QY 1 MISLPGVLTNLLRFLFLGLSALAPPRAQLQLP--ANRLQVSGGEVLPAYWTLHG 58
 DB 1 MILOQTPTSLLRVLFLGLSTLAAPRAQWELNVPGLNKLVAESGEVLPAYWTLAR 60
 QY 59 EVSSOPMEVFVWFFKQKEKE--DVLVSYNGVTTSKGVSILVYSPSRNLSLRLEGLO 117
 DB 61 EBSWSHREVLINFLPEQCKEPNQLVSYNGVMTNKGDTALVHSITSRNVSRLGALQ 120
 QY 118 EKSGPYSCSVNQDQKSGRHSIKTLELVLPAPPSCRLQGVPHVGNVLTSCSP 177